REMARKS

The amendments to the claims and the comments presented herein are responsive to both the Final Office Action dated December 30, 2003 (Paper No. 12222003) and the Advisory Action dated July 13, 2004 (Paper No. 070804). As indicated in the Advisory Action, the amendments to the claims presented in the Amendment and Response to Final Office Action dated May 28, 2004 have not been entered.

Claims 1-7 and 9-23 were pending in the application. Claims 1, 21, 22 and 23 have been amended and claim 9 has been canceled, without prejudice. Accordingly, after the amendments presented herein have been entered, claims 1-7 and 10-23 will remain pending.

Support for the amendments to the claims may be found through the specification including the originally filed claims. No new matter has been added.

Any amendments to and/or cancellation of the claims should in no way be constructed as an acquiescence to any of the Examiner's rejections and were done solely to expedite the prosecution of the application. Applicants reserve the right to pursue the claims as originally filed in this or a separate applications(s).

Response to Examiner's Comments Presented in the Advisory Action

In the Advisory Action the Examiner has indicated that "[t]he proposed amendment adds the limitation 'family of peptides that bind', which may raise [a] new matter rejection and the term is indefinite and further raise[s] new rejections under 35 USC. 112 first and second paragraphs."

Applicants respectfully submit that, in view of the teachings in Applicants' specification and the knowledge generally available in the art, the ordinarily skilled artisan would find the phrase "family of peptides that bind" to be clear and definite. To begin with, Applicants' specification discusses using the claimed methods to identify one or more compounds that bind to the luteinizing hormone releasing hormone receptor, a member of the G-protein coupled, seven transmembrane receptor *family of peptides* (see, for example, page 12, lines 24-27). Furthermore, the term "family of peptides" is well known in the art as indicated by, for example, J. K. Scott *et al.*, "A family of concanavalin A-binding peptides from a hexapeptide epitope

library", *Proc. Natl. Acad. Sci. USA* 89:5398-5402 (1992) and U.S. Patent No. 6,348,323 (Immunoassay of IGF *family of peptides,* their binding proteins and related molecules in dried whole blood filter paper spots), and U.S. Patent No. 6,706,493 (copies of which are attached herewith as Appendices A-C). The latter patent, in particular, claims an isolated DNA molecule "encoding a CCK receptor protein or a functional fragment thereof, whereby the fragment binds to the CCK receptor *family of peptides*" (see claim 1). Based on the foregoing, it is evident that the term "family of peptides" is well known in the art and intended to include a plurality of peptides that (by way of, for example, structural homology) are categorized as belonging to the same family.

In the Advisory Action the Examiner has also indicated that Applicants'arguments

> have been considered and are not persuasive, because applicants are referring to the reference non-peptide library to the first library as in the instant claims, the non-peptide library of the reference reads on the second library of the instant claims. Applicants arguments have been considered and are not persuasive for the following reasons: the reference gives guidance on how to make peptide libraries on solid supports, thus it would be obvious to one skilled in the art to make peptide libraries and use the libraries in further manipulations. Baindur et al teach active peptide identified (can be from a peptide library) through bioassay screening and can be optimized by synthesizing a large number of analogs by combinatorial parallel robotic synthesis. The reference further teaches that the active peptide that binds to the target is identified, and the active peptide is optimized synthesizing a large number of analogs, which method requires that the amino acid sequence of the active peptide is determined. Thus, the rejections of record have been maintained.

Applicants respectfully submit that, as admitted by the Examiner in the Final Office Action (see below) and as admitted by the Examiner in the Advisory Action (see above) Baindur *et al.* fail to teach or suggest methods which use a combination of a ***first peptide library*** and a ***second non-peptidic library***. In the Advisory Action, the Examiner appears to be saying that Applicants are referring to the Baindur *et al.* non-peptide library as the first library of the pending claims, whereas the non-peptide library of Baindur *et al.* corresponds to the second library of the instant claims. The Examiner is also generally asserting that "the reference gives guidance on how to

make peptide libraries on solid supports" and that, thus, "it would be obvious to one skilled in the art to make peptide libraries and use the libraries in further manipulations." To begin with, obviousness is not the standard to be used in a rejection under 35 U.S.C. §102(e). Moreover, as indicated in the response to the Final Office Action (and also re-iterated below), the Examiner has admitted that "[t]***he claimed invention differs from the prior art teaching by reciting 'forming a first library (peptide library)', and 'determining the amino acid sequence of at least one peptide that binds the target'***" and that the Baindur *et al.* reference "***does not specifically teach the use of the synthesized peptide library in further manipulations***." Thus, the pending claims are not anticipated under 35 U.S.C. §102(e) by Baindur *et al.* because this reference fails to teach each and every element of the claimed invention.

With respect to the rejection of the claims as being obvious over the teachings of Baindur *et al.*, Applicants respectfully submit that as indicated in the response to the Final Office Action (and also described in detail below) Baindur *et al.* fail to teach or suggest the claimed inventions and, thus, the teachings of Baindur *et al.* do not render the claimed invention obvious. Moreover, Baindur *et al.* teach away from the claimed invention in that they teach successful screening methods which employ peptide analogue-, peptidomimetic- and peptide derivative-based libraries as the source of the "lead peptide." Thus, the ordinarily skilled artisan reading Baindur *et al.* would not have been motivated to arrive at Applicants' invention, *i.e.*, use peptide libraries as the source of the lead peptide because Baindur *et al.* teach that the use of peptide analogue-, peptidomimetic- and peptide derivative-based libraries leads not only to the successful identification of leads, but also to leads that are more stable.

Finally, in the Advisory Action the Examiner has indicated that, with respect to the 35 U.S.C. §103(a) rejections over the combination of Hirshmann *et al.* and Blake or Hirshmann *et al.*, Blake and Gordon *et al.*, Applicants arguments have been considered and were not found to be persuasive because

the reference teaches 'design and synthesis of steroidal peptidomimetics, and mimics

> of peptides (refers to the peptide mimetics of the instant claims) are attached to the steroidal scaffold, which are designed based on the peptides which bind to the fibrinogen receptor (refers to the peptide), and the reference teaches at least steroid-peptide mimetics which refer to the library. Thus, applicants arguments are not persuasive. Applicants arguments are based on individual or single references whereas the rejection was based on combined teachings of Hirshmann et al and Blake et al. Blake et al teach methods of 'peptide library synthesis.' Thus it would have been obvious to one skilled in the art at the time the invention was made to use the peptides synthesized by the method taught by Blake et al and use the peptide as a model to design the peptide mimetics as taught by Hirshmann et al. And further applicants argue unexpected results, which are not present in the claims and further it would be obvious to one skilled in the art to obtain different results by optimizing the reaction conditions or reagents of the reference method. Applicants arguments regarding the rejection of claims over combined teachings of Hirshmann et al, Blake et al and Gordon et al have been considered and are not persuasive, since applicants arguments are based on individual references.

The Examiner is arguing that "the reference teaches 'design and synthesis of steroidal peptidomimetics, and mimics of peptides (refers to the peptide mimetics of the instant claims) are attached to the steroidal scaffold, which are designed based on the peptides which bind to the fibrinogen receptor (refers to the peptide)." The weakness of this argument lies in the fact that the four steroid RGD mimetics (1a-d) designed by Hirshmann *et al.* were not designed based on a peptide motif (or a peptide) that was generated by determining the amino acid sequence of a family of peptides selected from a library for binding to a target. The four steroid RGD mimetics of Hirshmann *et al.* were based on the peptides Echistatin and Kistrin (see page 9699, last line of the second column). Hirshmann *et al.* did not use a first library of peptides and did not select from such a library a peptide that binds to the target of interest. Thus, Hirshmann *et al.* fail to teach or suggest the claimed methods which require, *inter alia*, the foregoing steps.

With regard to the motivation of a skilled artisan to combine the teachings of Hirshmann *et al.* and Blake *et al.*, the Examiner has failed to address Applicants' arguments (set forth in detail below) indicating that "***[i]t is insufficient to establish obviousness that the separate elements of the invention existed in the prior art, absent some teaching or suggestion, in the prior art, to combine the elements***." *Arkie Lures v. Larew Tackle*, 119 F.3d. 953, (Fed. Cir. 1997).

Moreover, "[t]he mere fact that references can be combined or modified does not render the resultant combination obvious unless the prior art also suggests the desirability of the combination." M.P.E.P. § 2143.01. In the present case, not only do the references fail to teach or suggest the separate elements of the invention, but neither of the references contains any teaching or suggestion that would have motivated the skilled artisan to combine the teachings of Hirschmann *et al.* and Blake to arrive at the claimed invention. As indicated below in detail, Hirschmann *et al.* are concerned with a totally different problem, *i.e.*, studying the conformation of the peptides which bind to the fibrinogen receptor and, as a result, the skilled artisan looking to solve the problem solved by the inventors would not have even looked to Hirschmann *et al.* Similarly, Blake is devoid of any teaching that would have motivated the skilled artisan to look to non-peptide libraries. Thus, it is evident from the foregoing that that the Examiner has improperly relied on hindsight obtained from Applicants' invention in making the combination of references cited.

Acknowledgment of the Withdrawal of Certain Rejections

Applicants gratefully acknowledge the withdrawal of the following: (a) The previous rejection of claims 4, 8, 9 and 21 under 35 U.S.C §112, second paragraph; (b) The previous rejection of claim 21 under 35 U.S.C. 112, first paragraph; and (c) The previous obviousness type double patenting rejection over U.S. patent application serial no. 08/769,250.

Rejection of Claims 1-7 and 9-23 Under 35 U.S.C. §112, First Paragraph

The Examiner has rejected claims 1-7 and 9-23 under 35 U.S.C. §112, first paragraph, as failing to comply with the written description requirement. It is the Examiner's position that "[t]he claim(s) contains subject matter, which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention." In particular, the Examiner is of the opinion that

> [t]he specification description is directed to second library comprising analog library or the second library is synthesized based on altering D and L-amino acids; or the second library synthesized based on introduction of peptide mimetics at one or two positions within the library. The specification disclosure of [a] second library comprising non-peptide compounds clearly do[es] not provide an adequate representation regarding the open ended claimed non-peptide library compounds made and screened by the presently claimed invention. The instant[ly] claim[ed] non-peptide compounds would read on small organic molecule compounds [for] which [the] specification has no written support. ... In the present instance, the claimed invention contains no identifying characteristics regarding the non-peptide compounds of the second library. Additionally, the narrow scope of examples directed to peptide analogs or peptidomimetic compounds with non-natural amino acids are clearly not representative of the scope of non-peptide compounds of the presently claimed invention.

While in no way acquiescing to the validity of the Examiner's rejection and solely in the interest of expediting prosecution, Applicants have amended the claims to specify that the non-peptide compounds are peptide analogues, peptidomimetics or peptide derivatives, thereby rendering the foregoing rejection moot. Accordingly, Applicants respectfully request that the Examiner reconsider and withdraw the foregoing rejection.

Objection to Claim 9

The Examiner has objected to claim 9 under 37 CFR §1.75(c), "as being of improper dependent form for failing to further limit the subject matter of a previous claim" and has requested that Applicants cancel or amend the claim.

While in no way acquiescing to the validity of the Examiner's objection, Applicants have cancelled claim 9, thereby rendering the foregoing objection moot. Accordingly, Applicants respectfully request that the Examiner reconsider and withdraw the foregoing objection.

Rejection of Claims 1-7 and 9-23 nder §102(e)/§103 Over Baindur et al.

The Examiner has rejected claims 1-7 and 9-23 under 35 U.S.C. §102(e) as anticipated by or, in the alternative, under 35 U.S.C. §103(a) as obvious over Baindur et *al.* (U.S. Patent No. 5,891,737).

In particular, the Examiner is of the opinion that

> Baindur et al teach combinatorial non-peptide library . Baindur et al teach [that] peptide libraries are [the] source of small molecules having enormous structural diversity and can[*sic*] even larger conformational diversity. Active peptide[s] identified through bioassay screening (refers to instant claim step b)) can be quickly optimized by synthesizing a large number of analogs (refers to instant claim step d)) by combinatorial and/or parallel robotic synthesis (e.g., see column 3, lines 1-3). The reference teaches that peptide libraries generated using heterochiral amino acids, all D-amino acids and non-proteinogenic amino acids represent a rich source that can be mined for stabilized peptide leads. The reference teaches that these peptide leads are readily developed into non-peptide or peptidomimetic lead compounds (refers to instant[ly] claim[ed] non-peptide compounds). The reference teaches solid phase synthesis of peptide analogues or peptidomimetics. The reference teaches that as the number of variables within each building block (monomeric units, or chemical groups) increase, and/or as the number of building blocks increase, the size of [the] resultant library expands dramatically. Building blocks and monomers can be chemically conjugated to create a libraries containing components that are 100 to 1000 to 10,000 to 100,000 to 1,000,000 (and so on) building blocks in length (refers to instant claims 5-7).
> ***The claimed invention differs from the prior art teaching by reciting 'forming a first library (peptide library)', and 'determining the amino acid sequence of at least one peptide that binds the target'.***
> The reference teaches peptide libraries and methods of making the peptide libraries on solid support. ***The reference does not specifically teach the use of the synthesized peptide library in further manipulations***. However, the reference gives general guidelines to make solid phase synthesis of peptide libraries. Thus, it would have been obvious to one skilled in the art at the time the invention was filed how to make peptide libraries and use the libraries to obtain analogs. The reference specifically has not taught determining the amino acid sequence of the active peptide. However, the reference teaches that the active peptide that binds to the target is identified, and the active peptide is optimized synthesizing a large number of analogs. Thus, the reference would require the amino acid sequence of active peptides such that analogs of the peptides can be synthesized. The reference teaches the advantages of modified or analog peptides which are useful in therapy or diagnostics. Thus a person skilled in the art would have been motivated to make peptide libraries and use the active peptide from the library to prepare analogs which can be stabilized and useful in therapy. The claimed invention appears to be the same or obvious variations of the reference teachings, absent a showing of unobvious differences. (***Emphasis added***).

Applicants respectfully traverse the foregoing rejection on the grounds that Baindur et *al.* neither anticipate nor render obvious the claimed invention.

Rejection of Claims 1-7 and 9-23 Under 35 U.S.C.§102(e)

Applicants respectfully submit that Baindur et *al.* fail to teach the claimed invention for the following reasons.

The pending claims are directed to methods for identifying a non-peptide compound that binds to a target by (a) forming a first library comprising a multiplicity of peptides; (b) selecting from the first library a family of peptides that bind to the target; (c) determining the amino acid sequence or sequences of the family of peptides that bind to the target, thereby generating a peptide motif; (d) forming a second library comprising a multiplicity of non-peptide compounds designed based on the peptide motif, wherein said multiplicity of non-peptide compounds are selected from the group consisting of peptide analogues, peptidomimetics and peptide derivatives; (e) selecting from the second library at least one non-peptide compound that binds to the target; and (f) determining the structure or structures of the at least one non-peptide compound that binds to the target; thereby identifying a non-peptide compound that binds to the target.

For a prior art reference to anticipate in terms of 35 U.S.C. § 102 a claimed invention, the prior art must teach ***each and every element*** of the claimed invention. Lewmar Marine v. Barient, 827 F.2d 744, 3 USPQ2d 1766 (Fed. Cir. 1987).

Baindur et *al.* teach that

> [a]ctive peptides identified through bioassay screening ("lead peptides") can be quickly optimized by synthesizing a large number of analogs by combinatorial and/or parallel robotic synthesis. ***Peptide libraries generated using heterochiral amino acids, all D-amino acids and non-proteinogenic amino acids represent a rich source that can be mined for stabilized peptide leads***. These peptide leads can be used to further identify peptidomimetic or non peptide lead compounds. ***Libraries made using heterochiral amino acids or cyclic peptides can also generate peptide leads*** which are enzymatically stable and have constrained conformation(s). These types of peptide leads also are readily developed into non-peptide or peptidomimetic lead compounds.
> In the current era of drug development, high throughput screening of thousands to millions of compounds plays a key role. High throughput screening generally incorporates automation and robotics to enable testing these thousands to millions

> of compounds in one or more bioassays in a relatively short period of time. This high capacity screening technique requires enormous amounts of "raw materials" having immense molecular diversity to fill available capacity. Accordingly, combinatorial chemistry will play a significant role in meeting this demand for new molecules for screening. ***Once "leads" are identified using high throughput screening techniques, combinatorial chemistry will be advantageously used to optimize these initial leads*** (which analogs/variants will be tested in the same high throughput screening assay(s) that identified the initial lead). Thus, there is a need for new reagents and methods that can expand the scope of structural and conformational diversity contained within combinatorial libraries. (Column 3, lines 1-32; ***Emphasis added***).

Thus, Baindur et *al.* teach that the first library that leads to the identification of the "lead peptide" is composed of ***heterochiral amino acids, all D-amino acids and non-proteinogenic amino acids***. In contrast, Applicants' first library is composed of a multiplicity of ***peptides***. The term "peptide" is defined in the specification at page 4, lines 4 through 14 as follows:

> [t]he term "peptides," as used herein with regard to libraries, is intended to include ***molecules comprised only of natural amino acid residues*** (*i.e.*, alanine, arginine, aspeartic acid, asparagines, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenyalanine, praline, serine, threonine, tryptophan, tyrosine and valine) linked by peptide bonds, or other residues whose structures can be determined by standard sequencing methodologies (*e.g.*, direct sequencing of the amino acids making up the peptides or sequencing of nucleic acid molecules encoding the peptides). ***The term "peptide" is not intended to include molecules structurally related to peptides, such as peptide derivatives, peptide analogues or peptidomimetics***, whose structures cannot be determined by standard sequencing methodologies but rather must be determined by more complex chemical strategies, such as mass spectrometric methods. (***Emphasis added***).

In view of the foregoing it is evident that the methods of Baindur et *al.* differ from the claimed methods in that, *inter alia*, they use different types of libraries. In fact, the Examiner has admitted that "[t]he claimed invention differs from the prior art teaching by reciting 'forming a first library (peptide library)'."

Moreover, Baindur et *al.* fail to teach that ***a family of peptides that bind to the target*** is

selected from the first library and that the amino acid sequence or sequences of the family of peptides that bind to the target is determined, thereby generating a peptide motif upon which the second library is designed.

Since Baindur et *al.* fail to teach each and every element of the claimed invention, this reference does not anticipate the pending claims. Accordingly, the Examiner is respectfully requested to reconsider and withdraw this section 102(e) rejection.

Rejection of Claims 1-7 and 9-23 Under 35 U.S.C.§103

Moreover, Applicants respectfully traverse the Examiner's assertion that Baindur et *al.* would render the claimed invention obvious to the ordinarily skilled artisan at the time of the invention. Reconsideration and withdrawal of the rejection in light of the following discussion is respectfully requested.

As indicated above, the pending claims are directed to methods for identifying a non-peptide compound that binds to a target by (a) forming a first library comprising a multiplicity of peptides; (b) selecting from the first library a family of peptides that bind to the target; (c) determining the amino acid sequence or sequences of the family of peptides that bind to the target, thereby generating a peptide motif; (d) forming a second library comprising a multiplicity of non-peptide compounds designed based on the peptide motif, wherein said multiplicity of non-peptide compounds are selected from the group consisting of peptide analogues, peptidomimetics and peptide derivatives; (e) selecting from the second library at least one non-peptide compound that binds to the target; and (f) determining the structure or structures of the at least one non-peptide compound that binds to the target; thereby identifying a non-peptide compound that binds to the target.

To establish a *prima facie* case of obviousness, it is necessary for the Examiner to present evidence, preferably in the form of some teaching, suggestion, incentive or inference in the applied references, or in the form of generally available knowledge, that one having ordinary skill in the art would have been motivated to make the claimed invention and would have had a reasonable expectation of success in making the claimed invention. Under section 103, "[b]oth

the suggestion and the expectation of success must be founded in the prior art, not in applicant's disclosure" (*Amgen, Inc. v. Chugai Pharmaceutical Co., Ltd.* 927 F.2d 1200, 1207, 18 USPQ2d 1016 (Fed. Cir. 1991), quoting *In re Dow Chemical Co.,* 837 F.2d 469, 473, 5 USPQ2d 1529, 1531 (Fed Cir. 1988)).

Applicants submit that the Examiner has failed to establish a *prima facie* case of obviousness, since Baindur *et al.* fail to teach or suggest the claimed invention and further fail to provide the necessary motivation or expectation of success for the ordinarily skilled artisan to identify a non-peptide compound that binds to a target by forming a first library comprising ***a multiplicity of peptides***; selecting from the first library ***a family of peptides that bind to the target***; determining the sequence or sequences of the family of peptides that bind to the target, thereby generating a peptide motif; forming a second library comprising a multiplicity of peptide analogues, peptidomimetics and/or peptide derivatives designed based on the peptide motif; selecting from the second library at least one peptide analogue, peptidomimetic or peptide derivative that binds to the target; and determining the structure or structures of the at least one peptide analogue, peptidomimetic and/or peptide derivative that binds to the target, as required by Applicants' claims.

Baindur *et al.* fail to teach or suggest that the first library is composed of ***peptides*** and further fail to teach or suggest that ***a family of peptides that bind to the target*** is selected from the first library and that the amino acid sequence or sequences of the family of peptides that bind to the target is determined, thereby generating a peptide motif upon which the second library is designed.

In addition, Applicants respectfully submit that "[a] prior art reference must be considered in its entirety, i.e., as a whole, including portions that would lead away from the claimed invention. *W.L. Gore & Associates, Inc. v. Garlock, Inc.*, 721 F.2d 1540, 220 USPQ 303 (Fed. Cir. 1983), *cert. denied*, 469 U.S. 851 (1984)." M.P.E.P. § 2141.03. In the present case, Baindur *et al.* teach away from the claimed invention in that they teach successful screening methods which employ peptide analogue-, peptidomimetic- and peptide derivative-based libraries as the source of the "lead peptide." Thus, the ordinarily skilled artisan reading

Baindur *et al.* would not have been motivated to arrive at Applicants' invention, *i.e.*, use peptide libraries as the source of the lead peptide because Baindur *et al.* teach that the use of peptide analogue-, peptidomimetic- and peptide derivative-based libraries leads not only to the successful identification of peptide leads, but also to peptide leads that are more stable.

Based on the foregoing, it is evident that the Examiner has failed to establish a *prima facie* case of obviousness over Baindur *et al.* Accordingly, Applicants respectfully submit that this section 103(a) rejection is improper and request that it be reconsidered and withdrawn.

Rejection of Claims 1-23 Under 35 U.S.C. §112, Second Paragraph

The Examiner has maintained the rejection of claims 1-23 under 35 U.S.C. §112, second paragraph as being indefinite for the reasons of record. In particular, it is the Examiner's position that

> Applicant's arguments filed on 9/19/03, regarding 'non-peptide compounds as indefinite' have been fully considered but they are not persuasive. Applicant's traverse the rejection and submit that the term 'non-peptide' is clear and definite when read in light of specification. Applicant's arguments have been considered and are not persuasive. The 'non-peptide' as in applicant's specification disclosure could only read on peptide derivatives or analogues wherein a single amino acid or few specific amino acids are replaced by synthetic or non-natural amino acids. However, the specification has not disclosed how many amino acids replaced and which amino acids were replaced, and further the term 'non-peptide' may read on small organic molecules which were not the peptide derivatives or peptide analogues or contain the non-natural amino acids, as in applicants disclosure. Thus, the metes and bounds of the term 'non-peptide' is not clear.

While in no way acquiescing to the validity of the Examiner's rejection and solely in the interest of expediting prosecution, Applicants have amended the claims to specify that the non-peptide compounds are peptide analogues, peptidomimetics or peptide derivatives, thereby rendering the foregoing rejection moot. Accordingly, Applicants respectfully request that the Examiner reconsider and withdraw the foregoing rejection.

Rejection of Claims 1, 3-4 and 9-12 Under 35 U.S.C. §103(a) as Being Unpatentable Over Hirshmann et al. and Blake

The Examiner has maintained the rejection of claims 1, 3-4 and 9-12 under 35 U.S.C. §103(a) as being unpatentable over Hirshmann *et al.* and Blake for the reasons of record. In particular, it is the Examiner's position that

> Applicant's arguments have been considered and are not persuasive because Hirshmann et al teach non-peptide libraries of peptidomimetics based on known peptide motif. The rejection of record was based on [the]combined teachings of Hirshmann et al and Blake et al. Applicant's arguments against the references individually, one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. See/fl *re Keller,* 642 F.2d 413, 208 USPQ 871 (CCPA 1981); *In re Merck & Co.,* 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986). Hirshmann et al teach methods of making non-peptide libraries based on peptide motif, and methods of identifying the non-peptide which binds a target, and Blake teach methods of making peptide libraries and identifying a peptide which binds the target. Thus, it was well known in the art at the time the invention was made to methods for synthesizing peptide libraries and identifying a peptide which binds to a target, and the use the known peptide motif which binds to a target as a motif for the non-peptide libraries.
> Applicant's further argue that Blake do not teach 'second library.' Applicant's arguments against the references individually, one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. See *In re Keller,* 642 F.2d 413, 208 USPQ 871 (CCPA 1981); *In re Merck & Co.,* 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986). As discussed supra, in view of the combined teachings of Hirshmann et al and Blake, it would have been obvious to one skilled in the art at the time the invention was made to use the methods for making peptide libraries and screening the libraries for peptide (or a lead peptide) which binds to a target, and use the peptide in further non-peptide library synthesis. In response to applicant's argument that there is no suggestion to combine the references, the examiner recognizes that obviousness can only be established by combining or modifying the teachings of the prior art to produce the claimed invention where there is some teaching, suggestion, or motivation to do so found either in the references themselves or in the knowledge generally available to one of ordinary skill in the art. See *In re Fine,* 837 F.2d 1071, 5 USPQ2d 1596 (Fed. Cir. *I988)and In re Jones,* 958 F.2d 347, 21 USPQ2d 1941 (Fed. Cir. 1992). In this case, Hirshmann et al teach non-peptide libraries based on a known peptide ligand; and Blake teach methods for synthesizing a peptide libraries, and identifying a peptide which binds to a target. Hirshmann et al teach methods for making non-peptide libraries based on peptide motif, which peptide motif could be a peptide identified from a peptide library, and further at the time the invention was made it was well known in the art methods for synthesizing peptide libraries and identifying a peptide which binds to a target (Blake). Thus it would have been obvious to one skilled in the art methods of identifying a

peptide which binds to a target from a peptide library and further use the identified peptide in further libraries.
In response to applicant's argument that *'use of the claimed invention as described in the specification, allows for the identification of compounds that bind to a target that by use of peptide libraries with the use of the chemically based libraries such that advantage of each are maintained while the disadvantages of using either approach are overcome. The claimed invention has unexpectedly superior properties over the prior art because the skilled artisan can identify compounds that bind to a target by use of both peptide based and chemically based libraries, while maintaining the advantages of each,'* the fact that applicant has recognized another advantage which would flow naturally from following the suggestion of the prior art cannot be the basis for patentability when the differences would otherwise be obvious. See *Ex parte Obiaya, 221* USPQ 58, 60 (Bd. Pat. App. & Inter. 1985). Further it is not clear what are the advantages applicants are referring to which are different from the Hirshmann et al teachings. Hirshmann et al teach the known peptide motif which binds to a target and non-peptides which bind the target, thus the reference has compounds (peptide and non-peptide) which bind to the target. Applicant's arguments are not persuasive and the art rejections of record have been maintained for the reasons set forth in the previous office action.

The Combination of Hirshmann *et al.* and Blake Fail to Teach or Suggest the Claimed Invention

Applicants respectfully traverse the foregoing rejection on the grounds that the combination of Hirshmann *et al.* and Blake fail to teach or suggest the claimed invention. The Examiner insists that "Hirshmann et al teach methods of making non-peptide libraries based on [a] peptide motif, and methods of identifying the non-peptide which binds a target" (see above). However, Applicants fail to see where in Hirshmann *et al.* this teaching exists. ***The words "library," "motif," or "target" never appear in Hirshmann et al.*** Hirshmann *et al.* is merely a paper describing the attempt of the authors to study the bioactive conformation of the peptides which bind to the fibrinogen receptor. To do this, the authors designed four steroid RGD mimetics (1a-d) and tested them in a fibrinogen receptor assay. ***The design of four steroid mimetics for the purpose of studying the conformation of the peptides which bind to the fibrinogen receptor*** hardly amounts to a teaching or suggestion of methods for identifying a non-peptide compound that binds to a target by (a) forming a first library comprising a multiplicity of peptides; (b) selecting from the first library a family of peptides that bind to the target; (c) determining the amino acid sequence or sequences of family of peptides that bind to the target, thereby generating a peptide motif; (d) forming a second library

comprising a multiplicity of non-peptide compounds designed based on the peptide motif, wherein said multiplicity of non-peptide compounds are selected from the group consisting of peptide analogues, peptidomimetics and peptide derivatives; (e) selecting from the second library at least one non-peptide compound that binds to the target; and (f) determining the structure or structures of the at least one non-peptide compound that binds to the target, as required by Applicants' claims.

Moreover, for an Examiner to rely on a reference under 35 U.S.C.§103, the reference must be analogous prior art. (See M.P.E.P. 2141.01(a)). A prior art reference is analogous if the reference is in the field of applicant's endeavor or, if not, the reference is reasonably pertinent to the particular problem with which the inventor was concerned. *In re Oetiker*, 977 F.2d 1443, 1446, 24 USPQ2d 1443, 1445 (Fed. Cir. 1992). A reference is reasonably pertinent if, even though it may be in a different field from that of the inventor's endeavor, it is one which, because of the matter with which it deals, logically would have commended itself to an inventor's attention in considering his problem. *Wang Laboratories Inc. v. Toshiba Corp.*, 993 F.2d 858, 26 USPQ2d 1767 (Fed. Cir. 1993). M.P.E.P. §2145. In the present case, Hirschmann *et al.* is in the field of studying the bioactive conformation of the peptides which bind the fibrinogen receptor. As such, Hirschmann *et al.* is not in the field of Applicants' endeavor, *i.e.*, the field of high throughput library screening for identifying compounds that bind to a target. More importantly, Hirschmann *et al.* is not reasonably pertinent to the problem with which the inventors were concerned because a person having ordinary skill in the art would not reasonably have expected to solve the problem of successfully screening libraries containing a vast number of compounds to identify a compound that binds to a target, by considering a reference dealing with the bioactive conformation of the peptides which bind the fibrinogen receptor. Accordingly, Applicants respectfully submit that Hirschmann *et al.* is non-analogous prior art (as defined in M.P.E.P. §2141.01(a)) and, as such is not relevant to the obviousness of the claimed invention since the skilled artisan would not have looked to this article in considering the problem addressed by the inventors.

Furthermore, the secondary reference of Blake does not make up for the deficiencies of the primary reference. As indicated in the previous Amendment and Response, Blake teaches "methods for determining the amino acid sequence of peptides which bind to a ligand of interest from a large mixture of random or semi-random peptides" (column 2, lines 33-36). Blake does not teach or suggest forming a ***second library*** comprising a multiplicity of ***non-peptide compounds designed based on the peptide motif***; selecting from the second library at least one non-peptide compound that binds to the target; and determining the structure or structures of the at least one non-peptide compound that binds to the target, as required by Applicants' claims.

In response to the Examiner's assertion that "one cannot show nonobviousness by attacking references individually," Applicants respectfully submit that Applicants are not attacking the references individually. Applicants are addressing the combined teachings of Hirschmann *et al.* and Blake, however, for the sake of clarity Applicants present their arguments for each reference in a separate paragraph. As part of the analysis of the combined teachings of Hirschmann *et al.* and Blake, Applicants respectfully submit that "***[i]t is insufficient to establish obviousness that the separate elements of the invention existed in the prior art, absent some teaching or suggestion, in the prior art, to combine the elements***." *Arkie Lures v. Larew Tackle*, 119 F.3d. 953, (Fed. Cir. 1997). Moreover, "[t]he mere fact that references can be combined or modified does not render the resultant combination obvious unless the prior art also suggests the desirability of the combination." M.P.E.P. § 2143.01. In the present case, not only do the references fail to teach or suggest the separate elements of the invention, but neither of the references contains any teaching or suggestion that would have motivated the skilled artisan to combine the teachings of Hirschmann *et al.* and Blake to arrive at the claimed invention. As indicated above, Hirschmann *et al.* are concerned with a totally different problem, *i.e.*, studying the conformation of the peptides which bind to the fibrinogen receptor and, as a result, the skilled artisan looking to solve the problem solved by the inventors would not have even looked to Hirschmann *et al.* Similarly, Blake is devoid of any teaching that would have motivated the skilled artisan to look to non-peptide libraries. Thus, it is evident from the foregoing that that the

Examiner has improperly relied on hindsight obtained from Applicants' invention in making the combination of references cited.

Notwithstanding all of the foregoing and in the interest of expediting prosecution, Applicants have amended the claims, thereby further distinguishing the claimed invention from the cited prior art. Specifically, the claims are now directed to methods which involve ***selecting from the first library a family of peptides that bind to the target*** and determining the amino acid sequence or sequences of the family of peptides that bind to the target, thereby generating the peptide motif that is used to prepare the second non-peptide library. Applicants respectfully submit that for the reasons set forth above neither Hirschmann *et al.* nor Blake, alone or in combination, teach or suggest such methods.

Applicants' Unexpected Results Further Demonstrate That the Examiner Has Failed to Establish a *Prima Facie* Case of Obviousness

Even assuming *arguendo* that a *prima facie* case of obviousness were established by the Examiner, which Applicants dispute, the non-obviousness of the invention is apparent from the results achieved when the invention is put into practice. "One way for an Applicant to rebut a *prima facie* case of obviousness is to make a showing of 'unexpected results', *i.e.*, to show that the claimed invention exhibits some superior property or advantage that a person of ordinary skill in the relevant art would have found surprising or unexpected." In re Soni, 54 F.23d 746, 34 USPQ2d 1684 (Fed. Cir. 1995).

Use of the claimed invention as described in the specification, allows for the identification of compounds that have unexpectedly improved binding properties. Applicants refer the Examiner to PCT Publication WO97/22617 (submitted herewith as Appendix D), which is a published international application corresponding to the claimed invention. Example 2, at pages 19-26 of this PCT application, presents data demonstrating that relative to the unselected library, the potency (in terms of binding affinity for a substrate) of the selected population of non-peptide compounds (selected according to the claimed methods) was increased by

approximately ***1000-fold*** (see in particular page 24, lines 30-32). It is, therefore, apparent that the actual results obtained through use of the instantly claimed invention are not predicted from (*i.e.*, are unexpected over) the teachings of the prior art.

Based on all of the foregoing, it is evident that the Examiner has failed to establish a *prima facie* case of obviousness. Accordingly, Applicants respectfully submit that this section 103(a) rejection is improper and request that it be reconsidered and withdrawn.

Rejection of Claims 1, 3-7, 9-15 and 18-20 Under 35 U.S.C. §103(a) as Being Unpatentable Over Hirshmann et al., Blake and Gordon et al.

The Examiner has maintained the rejection of claims 1, 3-7, 9-15 and 18-20 under 35 U.S.C. §103(a) as being unpatentable over Hirshmann *et al.*, Blake and Gordon *et al.* for the reasons of record. In particular, it is the Examiner's position that

> Applicant's arguments filed on 9/19/03, regarding the rejection of claims over Hirshmann et al, Blake, and Gordon et al, have been fully considered but they are not persuasive. Applicants argue that Hirshmann et al and Blake fail to teach the claimed method. Applicant's arguments have been considered and are not persuasive. Applicant's arguments regarding Hirshmann et al and Blake have been addressed supra.. Applicants further argue that Gordon et al teach methods for making peptide libraries and screening strategies, and Gordon et al do not teach or suggest methods for forming second libraries comprising non-peptide libraries based on peptide motif that is identified by screening a primary library. In response to applicant's arguments against the references individually, one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. See/w *re Keller,* 642 F.2d 413, 208 USPQ 871 (CCPA 1981); *In re Merck & Co.,* 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986). Thus, it would have been obvious to one skilled in the art at the invention was made to form peptide libraries, and screen the library for a peptide which binds to a target as taught by Blake and use the known peptide(could be derived from a library) as basis for non-peptide libraries and screening them as taught by Hirshmann et al, and Gordon et al teach methods for making the specific size of the libraries.

Applicants disagree that the claimed invention would have been obvious to the ordinarily skilled artisan at the time it was made for the following reasons. First, as demonstrated above,

the combination of Hirschmann *et al.* and Blake, cited by the Examiner, fails to teach or suggest the claimed invention.

Moreover, the Gordon *et al.* reference fails to make up for the aforementioned deficiencies in the Hirschmann *et al.* and the Blake references and, thus, ***the combination of Hirschmann et al., Blake and Gordon et al. fail to teach or suggest the claimed invention***. Gordon *et al.*, teach methods for creating ***peptide*** combinatorial libraries and library screening strategies. Gordon *et al.* do not teach or suggest methods for forming a ***second library*** comprising a multiplicity of ***non-peptide compounds designed based on a peptide motif that is identified by screening a primary library***, as required by Applicants' claims. Rather, Gordon *et al.* only disclose, in a very general manner, that initial peptide leads identified in random library screening, "can serve as starting points for creating secondary recombinant ***peptide*** libraries or as leads for refinement by synthetic chemical combinatorial methods" (see Gordon *et al.*, page 1395, 1st column, lines 3-6). This statement fails to provide the necessary motivation for the ordinarily skilled artisan to make a ***non-peptidic secondary library based on peptide motif identified by screening of a primary peptide library***, as required by Applicants' claims.

Rather than addressing Applicants' arguments presented in the Amendment and Response dated September 19, 2003, the Examiner merely states that "one cannot show nonobviousness by attacking references individually." As indicated above, Applicants are not attacking the references individually. Rather, Applicants are addressing the combined teachings of Hirschmann *et al.*, Blake and Gordon *et al.* and respectfully submit that the combined teachings of these references fail to lead to the claimed invention for the reasons set forth above. Moreover, the Examiner has failed to provide any reasoning as to why the skilled artisan would have been motivated to combine the teachings of these references when *nothing* in these references provides the necessary motivation for the ordinarily skilled artisan to do so.

In summary, Applicants respectfully submit that based on the combination of the cited references, the Examiner has failed to establish a *prima facie* case of obviousness. Accordingly, Applicants respectfully submit that this section 103(a) rejection is improper and request that it be reconsidered and withdrawn.

Rejection of Claims 1, 3- 4, 9-12 and 16-17 Under 35 U.S.C. §103(a) as Being Unpatentable Over Hirshmann et al., Blake and Stankova

The Examiner has maintained the rejection of claims 1, 3- 4, 9-12 and 16-17 under 35 U.S.C. §103(a) as being unpatentable over Hirshmann *et al.*, Blake and Stankova for the reasons of record. In particular, it is the Examiner's position that

> Applicant's arguments filed on 9/19/03, regarding the rejection of claims over Hirshmann et al, Blake and Stankova et al have been fully considered but they are not persuasive. Applicants argue that Hirshmann et al and Blake fail to teach the claimed method. Applicant's arguments have been considered and are not persuasive. Applicant's arguments regarding Hirshmann et al and Blake have been addressed supra.. Applicants further argue that Stankova et al disclose screening non-peptide libraries by mass spectroscopy but not disclosed forming a first library and selecting from the first library one peptide that binds to the target. This rejection was based on combined teachings of Hirshmann et al, Blake and Stankova et al. In response to applicant's arguments against the references individually, one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. See *In re Keller,* 642 F.2d 413, 208 USPQ 871 (CCPA 1981); *In re Merck &* Co., 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986). Stankova et al teach the use of mass spectroscopy in analysis of structure of compounds identified in a library; Hirshmann et al teach methods of making non-peptide libraries and screening the libraries based on a known peptide motif(which binds to a target), and Blake teach the methods of making peptide libraries and methods of screening for a peptide which binds to a target. Thus, it would have been obvious to one skilled in the art at the invention was made to form peptide libraries, and screen the library for a peptide which binds to a target as taught by Blake and use the known peptide(could be derived from a library) as basis for non-peptide libraries and screening them as taught by Hirshmann et al and use tandem mass spectroscopy in identifying the structure of the non-peptide compounds as taught by Stankova et al.

Applicants disagree that the claimed invention would have been obvious to the ordinarily skilled artisan at the time it was made for the following reasons. First, as demonstrated above, the combination of Hirschmann *et al.* and Blake, cited by the Examiner, fails to teach or suggest the claimed invention.

Furthermore, Stankova *et al.*, relied on by the Examiner, does not make up for the above stated deficiencies in the Hirschmann *et al.* and the Blake references and, thus, ***the combination***

of Hirschmann et al., Blake and Stankova et al. fail to teach or suggest the claimed invention. Stankova *et al.* disclose screening non-peptide libraries by mass spectroscopy but do not disclose forming a first library comprising a multiplicity of peptides; selecting from the first library ***a family of peptides*** that bind to the target; determining the sequence or sequences of the family of peptides that bind to the target, thereby generating a peptide motif; forming a ***second library*** comprising a multiplicity of ***non-peptide compounds designed based on the peptide motif***; selecting from the second library at least one non-peptide compound that binds to the target; and determining the structure or structures of the at least one non-peptide compound that binds to the target, as required by Applicants' claims.

Rather than addressing Applicants' arguments presented in the Amendment and Response dated September 19, 2003, the Examiner merely states that "one cannot show nonobviousness by attacking references individually." As indicated above, Applicants are not attacking the references individually. Rather, Applicants are addressing the combined teachings of Hirschmann *et al.*, Blake and Stankova *et al.* and respectfully submit that the combined teachings of these references fail to lead to the claimed invention for the reasons set forth above. Moreover, the Examiner has failed to provide any reasoning as to why the skilled artisan would have been motivated to combine the teachings of these references when *nothing* in these references provides the necessary motivation for the ordinarily skilled artisan to do so.

In summary, Applicants respectfully submit that based on the combination of the cited references, the Examiner has failed to establish a *prima facie* case of obviousness. Accordingly, Applicants respectfully submit that this section 103(a) rejection is improper and request that it be reconsidered and withdrawn.

Rejection of Claims 1, 3-4, 8-12, 16-17 and 22-23Under 35 U.S.C. §103(a) as Being Unpatentable Over Hirshmann et al., Blake, Stankova and Scott et al.

The Examiner has maintained the rejection of claims 1, 3-4, 8-12, 16-17 and 22-23 under 35 U.S.C. §103(a) as being unpatentable over Hirshmann *et al.*, Blake, Stankova and Scott *et al.*

for the reasons of record. In particular, it is the Examiner's opinion that

> Applicant's arguments filed on 9/19/03, regarding the rejection of claims over Hirshmann et al, Blake and Stankova et al and Scott et al have been fully considered but they are not persuasive.
> Applicant's arguments have been considered and are not persuasive. Applicant's arguments regarding Hirshmann et al and Blake and Stankova et al have been addressed supra. Applicants further argue that Scott et al do not disclose forming a first library comprising multiplicity of peptides and use the peptide from the library in further methods of making non-peptide libraries. . Applicant's arguments have been considered and are not persuasive. Applicant's arguments against the references individually, one cannot show nonobviousness by attacking references individually where the rejections are based on combinations of references. See *In re Keller,* 642 F.2d 413, 208 USPQ 871 (CCPA 1981); *In re Merck & Co.,* 800 F.2d 1091, 231 USPQ 375 (Fed. Cir. 1986). The instant rejection was based on combined teachings of Hirshmann et al, Blake, Stankova et al and Scott et al. Hirshmann et al teach methods of making non-peptide libraries and screening the libraries based on a known peptide motif(which binds to a target), and Blake teach the methods of making peptide libraries and methods of screening for a peptide which binds to a target; Stankova et al teach the use of mass spectroscopy in analysis of structure of compounds identified in a library; and Scott et al teach methods of making peptide libraries on bacteriophages. Thus, it would have been obvious to one skilled in the art at the invention was made to form peptide libraries, and screen the library for a peptide which binds to a target as taught by Blake and use the known peptide(could be derived from a library) as basis for non-peptide libraries and screening them as taught by Hirshmann et al and use tandem mass spectroscopy in identifying the structure of the non-peptide compounds as taught by Stankova et al, and Scott et al the bacteriophage display peptide libraries.

Applicants disagree that the claimed invention would have been obvious to the ordinarily skilled artisan at the time it was made for the following reasons. First, as demonstrated above, the combination of Hirschmann *et al.*, Blake and Stankova, cited by the Examiner, fails to teach or suggest the claimed invention.

Furthermore, Scott *et al.*, relied on by the Examiner, does not make up for the above stated deficiencies in the primary references and, thus, ***the combination of Hirschmann et al., Blake, Stankova et al. and Scott et al. fail to teach or suggest the claimed invention***. Specifically, Scott *et al.* teach that the sequence of the peptide can be deduced from the determined DNA sequence of the carrier phage, but do not disclose forming a first library

comprising a multiplicity of peptides; selecting from the first library a family of peptides that bind to the target; determining the sequence or sequences of the family of peptides that bind to the target, thereby generating a peptide motif; forming a ***second library*** comprising a multiplicity of ***non-peptide compounds designed based on the peptide motif***; selecting from the second library at least one non-peptide compound that binds to the target; and determining the structure or structures of the at least one non-peptide compound that binds to the target.

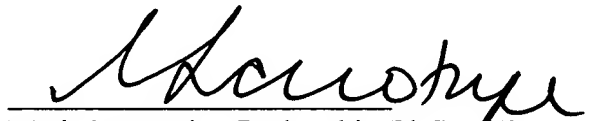
Rather than addressing Applicants' arguments presented in the Amendment and Response dated September 19, 2003, the Examiner merely states that "one cannot show nonobviousness by attacking references individually." As indicated above, Applicants are not attacking the references individually. Rather, Applicants are addressing the combined teachings of Hirschmann *et al.*, Blake, Stankova *et al.* and Scott *et al.* and respectfully submit that the combined teachings of these references fail to lead to the claimed invention for the reasons set forth above. Moreover, the Examiner has failed to provide any reasoning as to why the skilled artisan would have been motivated to combine the teachings of these references when *nothing* in these references provides the necessary motivation for the ordinarily skilled artisan to do so.

In summary, Applicants respectfully submit that based on the combination of the cited references, the Examiner has failed to establish a *prima facie* case of obviousness. Accordingly, Applicants respectfully submit that this section 103(a) rejection is improper and request that it be reconsidered and withdrawn.

CONCLUSION

In view of the amendments and remarks set forth above, it is respectfully submitted that this application is in condition for allowance. If there are any remaining issues or the Examiner believes that a telephone conversation with Applicants' Attorney could be helpful in expediting prosecution of this application, the Examiner is invited to call the undersigned at (617) 227-7400.

Respectfully submitted,



Maria Laccotripe Zacharakis, Ph.D., J.D.
Attorney for Applicants
Registration No. 56,266

LAHIVE & COCKFIELD, LLP
28 State Street
Boston, MA 02109
Tel. (617) 227-7400

Dated: December 3, 2004

Proc. Natl. Acad. Sci. USA
Vol. 89, pp. 5398–5402, June 1992
Biochemistry

A family of concanavalin A-binding peptides from a hexapeptide epitope library

(lectin/carbohydrate binding site/peptide mimics/peptide library/filamentous bacteriophage)

JAMIE K. SCOTT*†, DURAIKKANNU LOGANATHAN‡, R. BLAINE EASLEY*, XIAOFEN GONG*, AND IRWIN J. GOLDSTEIN‡

*Division of Biological Sciences, University of Missouri, Columbia, MO 65211; and ‡Department of Biological Chemistry, University of Michigan, Ann Arbor, MI 48109

Communicated by Stuart Kornfeld, March 12, 1992

ABSTRACT **The lectin concanavalin A (Con A) binds methyl α-D-mannopyranoside (MeαMan) as well as α-D-mannosyl groups at the nonreducing terminus of oligosaccharides. Ligand peptides that mimic the binding of MeαMan to Con A were identified from screening an epitope library composed of filamentous phage displaying random hexapeptides. A consensus sequence was identified among affinity-purified phage; Con A binds phage bearing this sequence and is inhibited from doing so by MeαMan. When tested for binding against a panel of lectins, phage bearing this sequence bind only weakly to a closely related D-mannose-binding lectin, indicating that binding to Con A is highly selective. A synthetic peptide bearing the consensus sequence blocks the precipitation of Con A by dextran with an inhibition strength equivalent to that of methyl α-D-glucopyranoside. These results demonstrate that the specificity of Con A is not limited to carbohydrates and that highly selective sugar-mimics for lectins of plant, animal, or bacterial origin may be identified from epitope libraries.**

The lectin Con A comprises a significant fraction of the protein in the jack bean (*Canavalia ensiformis*). It is a tetramer, consisting of four identical subunits that bind with moderate affinity (K_d 120–500 μM) to the α anomers of D-mannose and D-glucose (1). When oligosaccharides containing these sugars are displayed on a cell surface, Con A binds with high avidity as a result of multivalent interactions. Because of this high avidity and its specificity for particular sugars, the function of Con A is presumed to involve binding to D-mannosyl-containing oligosaccharides as part of a specific cell recognition process. In an effort toward developing specific target-ligands for lectins, we have discovered a set of related peptides that bind at or near the sugar-binding site of Con A with affinity equivalent to that of methyl α-D-glucopyranoside (MeαGlc). These peptides were obtained from screening a hexapeptide epitope library with Con A.

The epitope library consists of 200 million filamentous phage clones that display about 70% of all possible hexapeptides (2). The use of epitope libraries in identifying ligand peptides has been demonstrated for protein-binding antibodies (2–4), as well as for the biotin-binding site of streptavidin (5). The latter result was the first demonstration that epitope libraries can be used to discover peptides that mimic the binding of nonpeptide ligands.

Phage bearing Con A-binding sequences were affinity-enriched from the library and identified by (*i*) exhibiting a consensus sequence that is shared among randomly chosen phage and by their ability (*ii*) to selectively bind to Con A in ELISA, (*iii*) to be significantly retained by immobilized Con A, and (*iv*) to be specifically inhibited from binding and enrichment on Con A by methyl α-D-mannopyranoside (MeαMan). Furthermore, hexapeptide bearing the consensus sequence was shown to bind directly to Con A and to inhibit the precipitation of Con A by dextran with a strength equivalent to that of MeαGlc, whereas peptides bearing selected alanine replacements in the sequence inhibited binding to a lesser extent. Binding of the consensus sequence to Con A is selective, as phage bearing this sequence bind weakly to only one of several MeαMan-binding lectins that share structural homology with Con A and not at all to lectins that recognize other sugars. Further, the consensus sequence was not isolated in side-by-side screenings of the epitope library with Con A and several of these MeαMan-binding lectins.

The specificity of peptide recognition by Con A is demonstrated by the ability of the consensus sequence to directly bind Con A and to be inhibited from binding by MeαMan. In addition, the remarkable selectivity of this sequence for Con A is shown by its inability to bind strongly to other D-mannose-binding lectins. That we have identified peptides with these properties indicates that epitope libraries may be valuable in probing the specificity of both plant and animal lectins and more generally, in identifying lead compounds for development of receptor-targeted drugs.

MATERIALS AND METHODS

Isolation of Ligand Phage from the Hexapeptide Epitope Library. Phage from the hexapeptide library were isolated by "biopanning" (2, 6). About 10^{12} phage particles were incubated overnight with biotinylated Con A (bio-Con A) at 2 μM in Tris-buffered saline (TBS) containing 10 nM $MnCl_2$ and 100 nM $CaCl_2$ and adjusted to pH 7.0–7.2 (TBS/MnCa); soluble epidermal growth factor receptor (a gift of M. Das, University of Pennsylvania) was also present at half the molar concentration of the Con A. The reactions mixtures were incubated for 20 min on streptavidin-coated polystyrene Petri dishes. Unbound phage were removed by extensive washing in TBS/MnCa containing 0.5% Tween 20 (TBS/MnCa/Tween), and the remaining phage were eluted with acid. The eluates were neutralized and used to infect cells.

The strategy used in the series of four biopannings with three different concentrations of bio-Con A (2, 0.2, and 0.02 μM) is shown below.

1st	2nd	3rd	4th	
2	2	2	2	Isolate clones
	0.2	0.2	0.2	Isolate clones
	0.02	0.02	0.02	Isolate clones

Scheme I

Abbreviations: bio-Con A, biotinylated Con A; MeαGlc, methyl α-D-glucopyranoside; MeαMan, methyl α-D-mannopyranoside; NphαMan, *p*-nitrophenyl α-D-mannopyranoside.

†To whom reprint requests should be sent at present address: c/o Dr. Elizabeth Getzoff, The Scripps Research Institute, 10666 North Torrey Pines Road, La Jolla, CA 92037.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

A high concentration of bio-Con A (2 μM) was used in the first round of biopanning to maximize the yield of ligand phage. Three samples of amplified phage from the first round of biopanning were subjected to a second round of biopanning at 2, 0.2, or 0.02 μM bio-Con A, followed by amplification of the enriched phage. Two more rounds of biopanning at the three lectin concentrations and amplification were carried out in series, for a total of four rounds of biopanning. Clones from the fourth round of biopanning at each concentration of bio-Con A were chosen at random and propagated, and their DNA was sequenced in the epitope region (2).

Assessment of Phage Binding by Micropanning. Phage clones were propagated in *Escherichia coli* K91 cells and partially purified by two precipitations with polyethylene glycol (2), followed by pelleting by centrifugation at 58,000 rpm for 70 min at 15°C in a Beckman TLA rotor and resuspension in 1 ml of TBS/0.02% NaN_3. Virion concentrations were estimated from agarose gel electrophoresis (7). Lectins used in micropanning assays include bio-Con A and the biotinylated forms of *Lens culinaris* agglutinin, *Pisum sativum* agglutinin, *Sophora japonica* agglutinin, and *Griffonia simplicifolia* lectin (Vector Laboratories). Polystyrene microtiter wells were coated with 1 μg of streptavidin in 0.1 M $NaHCO_3$ at 4°C overnight or at 37°C for 4 hr, washed with TBS, and then incubated with about 50 pmol of biotinylated lectin in TBS/MnCa at 4°C overnight or at room temperature for 6 hr. Wells were washed six times by filling with TBS/MnCa/Tween, then 30 μl of TBS/MnCa/Tween or 5 mM Meα Man in TBS/MnCa/Tween was added to each well and incubated for 30 min at 4°C. About 10^8 phage particles in 5 μl of TBS/MnCa/Tween were added to each well and incubated at 4°C for 4 hr. The incubation mixtures were removed and the wells were washed 10 times (5 min per wash) with TBS/MnCa/Tween at 4°C. Retained phage were eluted from the wells by incubation for 10 min in 30 μl of 0.1 M HCl, pH 2.2 with glycine, containing bovine serum albumin at 1 μg/ml and then were transferred to microcentrifuge tubes containing 3.6 μl of 0.5 M Tris base. Phage from the reactions and the eluates were titered on *E. coli* K91 cells (8). Percent yields were estimated from the number of transforming units in the eluates divided by those in the reactions.

Detection of Binding of Biotinylated Lectin to Phage by ELISA. Microtiter wells were coated overnight at 4°C with about 10^{11} CsCl-purified (6) virions (≈700 fmol equivalents of phage-borne peptide), 700 fmol of horse gamma globulin, or 1 mg of crystallized bovine serum albumin (Pentex, Kankakee, IL) in 30 μl of 0.1 M $NaHCO_3$ or were incubated with buffer alone. Wells were washed three times in TBS/Tween, and then 30 μl of 10 nM biotinylated lectin in TBS/MnCa/Tween with or without 20 mM Meα Man was added to appropriate wells and incubated overnight at 4°C. Plates were developed with an ABC Elite kit (Vector Laboratories). Wells were washed seven times with cold TBS/MnCa/Tween and once with cold TBS/MnCa, then 30 μl of ABC solution was added to each well and incubated at room temperature for 30 min. Wells were washed seven times with cold TBS/MnCa/Tween, and 30 μl of developing solution was added to each well and incubated at room temperature for 1–2 hr. Solute optical density in each well was measured at 405 and 490 nm by a microplate reader.

Synthetic Peptides. About 30 μg of hexapeptide was synthesized on each polyethylene pin (see ref. 9; a generous gift of H. M. Geysen, Chiron Mimotopes, Victoria, Australia). Pins were cleaned before each use (9), dried, and pretested by incubation in ABC solution and development as described below. Then they were cleaned again and tested for bio-Con A binding. First, the pins were blocked by incubation for 4 hr at room temperature in microtiter wells with 0.5% nonfat dry milk/0.1% Tween 20/TBS, then washed in TBS for 5 min, transferred to microtiter wells containing 130 μl of 10 nM bio-Con A/0.05% dialyzed bovine serum albumin/TBS/MnCa with or without 5 mM Meα Man, and incubated at 4°C for 4–16 hr. Pins were washed four times in TBS/MnCa/Tween, transferred to wells containing 130 μl of ABC solution and incubated at room temperature for 30–45 min. The pins were washed four times, transferred to wells containing developing solution, and incubated for 45–75 min. The pins were removed, and the optical density of each well was measured.

The peptides MYWYPY, MAWAPA, and VGRAFS (kindly provided by T. Leung and W. Mandecki, Abbott) were synthesized on a 430A peptide synthesizer (Applied Biosystems), purified by reversed-phase HPLC in a gradient of acetonitrile in 0.1% trifluoroacetic acid, and lyophilized. The sequences of purified peptides were confirmed by mass spectroscopic and amino acid analysis, and in select cases, peptide concentrations were verified by amino acid analysis.

The synthetic peptides, Meα Man, and Meα Glc were tested for their ability to inhibit the precipitation of Con A by dextran (10). Stock solutions of Meα Man (10 mM), Meα Glc (10 mM), MAWAPA (4.9 mM), and VGRAFS (5.4 mM) were prepared both in water and in 10% (vol/vol) dioxane, whereas MYWYPY (1.62 mM) was soluble only in 10% dioxane. Reaction mixtures (100 or 200 μl) contained 18 μg of Con A, 15 μg of dextran B-1355, and sugar or peptide inhibitor in 1 M NaCl/10 mM phosphate, pH 7.2; when present, dioxane was at a final concentration of 2%. Reaction mixtures were incubated at room temperature for 48 hr. After microcentrifugation for 10 min, the pellets were washed with buffer and centrifuged three times; protein content in pellets was measured by the method of Lowry.

The peptide VGRAFS was tested for its ability to compete with the chromogenic ligand *p*-nitrophenyl α-D-mannopyranoside (Nphα Man) for the sugar-binding site of Con A. For this assay, difference spectra (range, 400–290 nm) were recorded on a Varian Cary 219 dual-beam spectrophotometer at 25°C with partitioned (Yankeelov, Beckman Instruments Fullerton, CA) cuvettes (11). Briefly, to one side of two matched cuvettes was added 1 ml of 80 μM Con A binding sites in phosphate-buffered saline, pH 7.2; to the other side, 1 ml of Nphα Man in the same buffer was added. Baseline was recorded for the two cuvettes, and then the sample cuvette was mixed and the difference spectrum was recorded, noting the absorbance maximum (ΔA) at 317 nm. The reference cuvette was then mixed to restore the baseline. Concentrations of 20–240 μM Nphα Man mixed with Con A in buffer were used to obtain the value ΔA_m, the change in absorbance corresponding to complete protein saturation; then cuvettes containing Con A in one side and 200 μM Nphα Man plus 0.2–1.5 mM peptide or 50–600 μM Meα Man in the other side were mixed to obtain difference spectra. The inhibitor K_d values were calculated from absorbance data (11).

RESULTS

A Sequence Motif Is Shared Among Phage Isolated with bio-Con A. Phage clones were randomly isolated after the final round of each series of biopannings, and their epitope regions were sequenced. Table 1 shows these sequences arranged into four groups, A–D. All of the sequences in group A share a single motif, YPY. The frequency with which phage bearing the YPY motif, including those bearing the sequence MYWYPY, occur increases concomitantly with the decrease in the concentration of bio-Con A used in biopanning. This result indicates that these sequences were affinity-selected and that they are among the tightest-binding sequences available in this epitope library, as their occurrence increases with stringency of selection. The sequences flanking the random hexapeptide region may also be involved in the selection of YPY by Con A, as the consensus sequence

Table 1. Epitope sequences of phage isolated after four rounds of biopanning with bio-Con A

	2 μM	0.2 μM	0.02 μM
Group A	MYWYPY (2)	MYWYPY (8)	MYWYPY (11)
	IPWYPY (2)	IPWYPY (7)	
	LYWYPY (6)	LYWYPY	
	RIFYPY (2)	IAWYPY (3)	YTWYPY
	PIFYPY	IFWYPY	PYWYPY
	LPFYPY	VWWYPY (2)	
	FYWYPY		
	YVYYPY (5)		
	FYYYPY		
Group B	PLFVRY	TAFQLS	VGRAFS
	TYSATV	HRVGGT	VSWYIS
	WFSFMS (2)		
	ENGRKS		
	SSSGFW		
	VPGVSF		
	ARRYSR		
	HSSYFF		
Group C	RRHHHH (2)	CACRLK	
	-LHHHH		
	HWLVHH		
	HRHKHQ		
Group D	WPDMVR		RAAGIV
	FNAAVL		

Sequences are shown below the concentration bio-Con A used for the last three rounds of biopanning. Parenthetic numbers signify the number of clones identified with the stated sequence. All repeatedly isolated sequences are encoded by identical nucleotide sequences, indicating that they derive from a single clone.

occurs in only one register, at the last three residues of the hexapeptide variable region. Moreover, an aromatic residue (W, F, or Y) always precedes the YPY consensus sequence and is most likely also involved in the selection process. To confirm our results, this biopanning experiment was repeated with minor changes (including the absence of soluble epidermal growth factor receptor). Once again, the YPY motif was observed in 11 of 20 clones, with 17 of the 20 having YXY and 15 of these 17 having an aromatic residue preceding the YXY sequence (data not shown). Furthermore, side-by-side biopannings with two other Meα Man-binding lectins (*L. culinaris* agglutinin and *P. sativum* agglutinin) that accompanied the repeat Con A biopanning failed to select this or any sequence motif (data not shown).

In loose accordance with group A, each sequence in group B includes at least one hydroxylated amino acid (S, T, or Y) as well as an aromatic residue. Group C comprises residues that could potentially interact with the transition metal that is located in the sugar-binding site: these are H and C. Interestingly, all of the group C sequences can potentially achieve stable interactions with Con A by forming bidentate chelates with Mn^{2+} (12). Group D includes the remaining sequences that were identified. The sequence data indicate that epitopes in groups B and C probably bind Con A, but with lower affinities than those in group A, since none of these sequences share a strict consensus, and their relative frequencies of occurrence decline as the stringency of selection increases among the parallel biopannings.

Phage-Borne Epitopes Bind Con A. Representative phage selected from groups A–C were tested by two assays for direct binding to Con A. Binding of bio-Con A to CsCl-purified, plate-immobilized phage was tested by ELISA (Table 2). Phage bearing WYPY bound bio-Con A well above background, as did those displaying RRHHHH. Phage-borne VGRAFS also bound lectin above background, but to a lesser degree. Binding to phage—as well as background binding to CsCl-purified f1 .phage, to albumin-coated wells, and to uncoated wells—was inhibited in the presence of Meα Man. Direct binding of affinity-selected phage to plate-immobilized Con A was also tested by micropanning assays (Table 2) in which the yield of phage obtained in a miniaturized biopanning procedure was compared with that of control phage bearing a non-binding sequence. Phage bearing the WYPY sequence motif exhibited high yields on micropanning, with MYWYPY producing the highest yield, indicating that phage bearing this sequence bind most tightly to Con A. Each of the affinity-selected phage clones tested bound Con A with some specificity, as the phage yield from all these clones was significantly decreased when panning was done in the presence of Meα Man. In contrast, the yield of phage bearing the control sequence DFLEKI increased under this condition.

Synthetic Peptides Bind Con A. The peptides MYWYPY, LYWYPY, and IAWYPY were each synthesized on two pins and assayed for binding to Con A in the absence or presence of 5 mM Meα Man. The pins were cleaned and pretested for background binding, then tested for their ability to bind Con A (Table 3). The peptide MYWYPY gave the strongest signal, with the other WYPY-bearing peptides giving relatively high signals compared with background. This binding was specific, as it was significantly decreased in the presence of Meα Man.

To determine relative affinity compared with Meα Man, the hexapeptides MYWYPY (from group A), MAWAPA (a control peptide), and VGRAFS (from group B) were synthesized

Table 2. Binding of bio-Con A to affinity-purified phage

	ELISA with immobilized phage*		Micropanning with immobilized bio-Con A†			
				Output TU		
Epitope sequence	No Meα Man	20 mM Meα Man	Input TU	No Meα Man	5 mM Meα Man	% inhibition
MYWYPY	1015	39	4.1×10^7	2.6×10^7	1.1×10^5	100
LYWYPY	1067	38	6.0×10^7	5.1×10^6	7.0×10^5	86
IAWYPY	1039	62	4.0×10^7	5.6×10^6	5.4×10^5	90
VGRAFS	829	29	8.2×10^7	6500	510	92
RRHHHH	1000	37	1.6×10^7	4.4×10^4	2600	94
DFLEKI	—	—	7.5×10^6	1900	6000	−220
f1 phage	529	25	—	—	—	—
Horse γG	1412	84	—	—	—	—
Albumin	683	18	—	—	—	—
Blank	585	18	—	—	—	—

Horse γG, horse gamma globulin; albumin, bovine serum albumin.

*Data were calculated as the average of two replicates [$(OD_{405} - OD_{490}) \times 10^3$].

†TU, transforming units. Percent inhibition was calculated as [(no Meα Man TU − 5 mM Meα Man TU)/no Meα Man TU] × 100.

Table 3. Binding of bio-Con A to peptides on pins

Epitope sequence	First pin		Second pin	
	Pretest	bio-Con A	Pretest	bio-Con A, MeαMan
MYWYPY	41	1069	46	325 (73)*
LYWYPY	43	682	46	204 (75)
IAWYPY	52	1108	52	444 (63)
Background*	—	43	—	44

The bio-Con A concentration was 10 nM, and MeαMan was 5 mM. Values are $(OD_{405} - OD_{490}) \times 10^3$. Percent inhibition was calculated as {1 − [bio-Con A, 5 mM MeαMan signal − pretest signal)/(bio-Con A signal − pretest signal)]} × 100 and is given in parentheses.
*Average of four wells that did not incubate a pin.

and tested for their ability to inhibit the precipitation of Con A by dextran (Fig. 1). The peptide MYWYPY inhibited precipitation with an IC_{50} of 490 μM, whereas the IC_{50} values of MeαMan and MeαGlc were 120 μM and 480 μM, respectively. The inhibitory strength of the control peptide MAWAPA was about half that of the parent sequence, indicating that tyrosine side chains contribute to binding. As expected, the group B peptide VGRAFS produced weaker inhibition that was similar to that of the control peptide MAWAPA.

The peptide VGRAFS was also tested for its ability to displace the chromagenic ligand NphαMan from the sugar-binding site of Con A. (Because of relative insolubility and insufficient quantity, the other peptides were not tested.) The plots of NphαMan displacement vs. competitor concentration are linear for VGRAFS and MeαMan (Fig. 2), indicating that both bind Con A competitively (11). The data produced K_d values of 80 μM for MeαMan and 1.62 mM for the peptide; these values are in rough agreement with the IC_{50} values calculated from the precipitation inhibition assays. By comparison, a K_d of 800 μM was estimated for the binding of MYWYPY to Con A.

Specificity of Phage Bearing Con A-Binding Epitopes. In micropanning experiments in which phage were panned on a panel of closely related MeαMan-binding lectins, including Con A, *L. culinaris* agglutinin and *P. sativum* agglutinin (13), the highest yields by far were obtained when phage bearing the WYPY motif were panned on Con A; a significant yield was also seen when phage bearing MYWYPY were panned on *P. sativum* agglutinin (Table 4). When immobilized phage were tested for binding to the above lectins by ELISA, again the strongest binding was seen with Con A, and weaker, but detectable binding was obtained for all of the peptides with pea lectin, but not with lentil lectin (Table 4). When *G. simplicifolia* I lectin and *S. japonica* agglutinin were tested by micropanning assays and ELISA, these lectins, which do not bind MeαMan, showed no binding to the phage (data not shown).

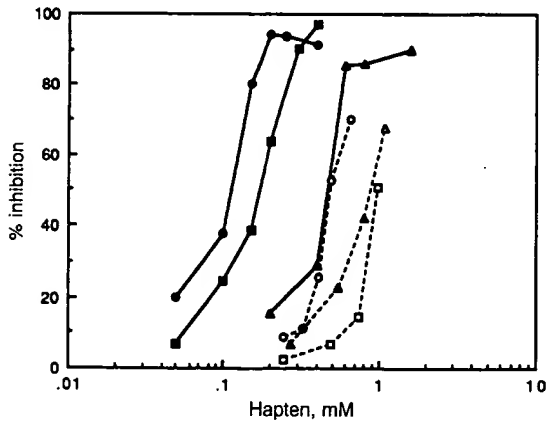


FIG. 1. Inhibition of the precipitation reaction between Con A and dextran B-1355-S by MeαMan (●), aqueous MeαMan (■), MeαGlc (▲), and the synthetic peptides MYWYPY (○), MAWAPA (□), and VGRAFS (△). IC_{50} values are as follows: MeαMan, 0.12 mM; aqueous MeαMan, 0.18 mM; MeαGlc, 0.48 mM; MYWYPY, 0.49 mM; MAWAPA, 0.98 mM; and VGRAFS, 0.9 mM. Vials contained 2% dioxane except for a control experiment done in the absence of dioxane (aqueous MeαMan).

DISCUSSION

This work demonstrates that two very different molecules, carbohydrate and peptide, are recognized by Con A with similar affinities and that, most likely, the ligand peptides described herein competitively inhibit Con A–MeαMan interaction by binding at or near the sugar-binding site. MYWYPY, the most promising peptide isolated from the hexapeptide epitope library, binds Con A with a K_d of about 800 μM, equivalent to that of MeαGlc, yet, surprisingly, is highly selective, as it is able to distinguish between closely related lectins that bind MeαMan.

Recent crystallographic studies of the structure of Con A in complex with MeαMan show that the sugar is bound in the C1 chair conformation via a network of seven hydrogen bonds that link oxygen atoms O-3, O-4, O-5, and O-6 of the sugar to residues Asn-14, Leu-99, Tyr-100, Asp-208, and Arg-228, as well as by hydrophobic interactions between C-5 and C-6 of the sugar and the ring atoms of Tyr-12 and Tyr-100 (14). It is tempting to speculate that epitopes bearing the WYPY motif bind as a result of both hydrogen bonding with the hydroxyl groups of tyrosine residues (which mimic sugar oxygens) and hydrophobic interactions with carbons on the aromatic side chain (which mimic sugar carbons). In support of this, replacement of tyrosine in the consensus sequence with alanine resulted in an ≈50% reduction in inhibitory strength (Fig. 1). Furthermore, all of the group B sequences have at least one residue with a hydroxylated side chain (Y, T, or S), and all but one have at least one aromatic residue. In keeping with this, the group B peptide we tested,

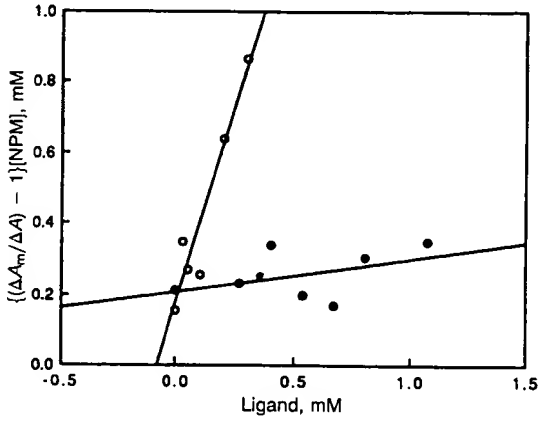


FIG. 2. Displacement of NphαMan from Con A by competing ligands (○, MeαMan; ●, peptide VGRAFS). ΔA_m is the change in absorbance corresponding to complete saturation of Con A with NphαMan, ΔA the change in absorbance for every addition of the competing ligand, and [NPM] the concentration of free NphαMan. Concentration of total NphαMan was 200 μM and that of Con A was 20 μM, corresponding to 40 μM combining sites. Dissociation constants, K_d, determined from this plot for MeαMan and VGRAFS are 0.08 mM and 1.62 mM, respectively.

Table 4. Binding of biotinylated lectins to affinity-purified phage

Epitope sequence	MeαMan	Micropanning with immobilized lectin*				ELISA with immobilized phage†		
		bio-Con A	bio-LCA‡	bio-PSA	Buffer	bio-Con A	bio-LCA	bio-PSA
MYWYPY	−	1733	2.0	60	0.7	982	23	50
	+	73	0.7	0.7	1.3	71	18	19
LYWYPY	−	600	3.0	1.0	1.0	845	20	52
	+	55	0.5	0.3	1.0	80	17	17
VGRAFS	−	1.4	1.2	2.0	0.2	574	19	42
	+	0.8	2.0	1.6	2.0	40	19	18
RRHHHH	−	2.3	3.0	0.3	1.0	840	27	68
	+	2.0	1.0	0.8	1.0	50	17	17
DFLEKI	−	<1.0	>1.0	2.5	<1.0	—	—	—
	+	<1.0	1.0	<1.0	1.0	—	—	—
f1 phage	−	—	—	—	—	352	21	27
	+	—	—	—	—	94	19	18
Fetuin	−	—	—	—	—	1561	562	435
	+	—	—	—	—	53	86	24

Biotinylated lectins included *L. culinaris* agglutinin (bio-LCA) and *P. sativum* agglutinin (bio-PSA).
*Micropanning data are relative yields and are calculated as percent yield obtained from micropanning phage on lectin-coated wells divided by the average percent yield from two wells that were treated with buffer alone.
†ELISA data are the average ($OD_{405} - OD_{490} \times 10^3$) from two wells.

VGRAFS, bound less tightly than MYWYPY but bound specifically to Con A (Figs. 1 and 2; Table 2).

Why do not other, closely related Meα Man-binding lectins recognize Con A-binding peptides? In general, the structure of the sugar-binding sites of these lectins is highly conserved, however, the sugar-contact residues Arg-228, Leu-99 and Tyr-100 of Con A are replaced by Gly-99, Ala-210 and Glu-211 in favin, pea, and lentil lectins (15). These differences may be involved in the observed binding specificity for the ligand peptides that Con A selected from the epitope library. However, the binding of peptide to Con A probably involves more than the residues known to contact MeαMan, as the consensus sequence is four residues long; these as yet undefined contacts may be responsible for the selective binding to Con A. In support of this, phenyl α-D-mannopyranoside is known to bind Con A, favin, pea, and lentil lectins with higher affinity than MeαMan, indicating that there are other conserved residues in the sugar-binding site that are available for hydrophobic interaction (13, 16). It should be possible to determine which residues contribute to binding by studying the crystallographic structure of peptide in complex with Con A (14).

The approach described here may be applied to the discovery of ligand peptides that specifically block the binding of biomedically significant animal lectins to their biological targets (e.g., the vascular E-selectin, ELAM-1, which is thought to target neutrophils and T cells to areas of inflammation; ref. 17). Although, in general, peptides themselves make poor drug candidates, they may serve as lead compounds, providing chemical information that would be used to develop pharmacologically active agents (18); selectivity of binding is a key characteristic of good lead compounds. Our results are particularly encouraging for the use of epitope libraries in identifying lead compounds, as the ligand peptides described here have high selectivity, albeit like their sugar counterparts, moderate affinity.

Many thanks to Tat Leung and Wlodek Mandecki for providing synthetic peptides and useful discussions, to H. Mario Geysen for providing pin-immobilized synthetic peptides, to Vincent Massey and Harry Winter for assistance in the displacement assay, to Cindy Schopp and Robert Davis for excellent technical assistance, and to George Smith, Rosalind Kornfeld, and Baldomero Olivera for helpful discussions. This work was supported by National Institutes of Health Grants GM41478 (G. P. Smith, University of Missouri-Columbia) and GM29470 (I.J.G.), National Research Service Award GM13772 (J.K.S.), Howard Hughes Medical Institute Undergraduate Internship (R.B.E.), a Sandoz Summer Fellowship (R.B.E.), a Protos contract (G. P. Smith), and by the Department of Medicine and Molecular Biology Program at the University of Missouri.

1. Goldstein, I. J., Hollerman, C. E. & Smith, E. E. (1965) *Biochemistry* **4**, 876–883.
2. Scott, J. K. & Smith, G. P. (1990) *Science* **249**, 386–390.
3. Cwirla, S. E., Peters, E. A., Barrett, R. W. & Dower, W. J. (1990) *Proc. Natl. Acad. Sci. USA* **87**, 6378–6382.
4. Felici, F., Castagnoli, L., Musacchio, A., Jappelli, R. & Cesareni, G. (1991) *J. Mol. Biol.*, **222**, 301–310.
5. Devlin, J. J., Panganiban, L. C. & Devlin, P. E. (1990) *Science* **249**, 404–406.
6. Smith, G. P. & Scott, J. K. (1992) *Methods Enzymol.* **217**, in press.
7. Smith, G. P. (1988) *Virology* **167**, 155–165.
8. Parmley, S. F. & Smith, G. P. (1988) *Gene* **73**, 305–318.
9. Geysen, H. M., Rodda, S. J., Mason, T. J., Tribbick, G. & Schoofs, P. G. (1987) *J. Immunol. Methods* **102**, 259–274.
10. So, L. L. & Goldstein, I. J. (1967) *J. Immunol.* **99**, 158–163.
11. Bessler, W., Shafer, J. A. & Goldstein, I. J. (1974) *J. Biol. Chem.* **249**, 2819–2822.
12. Arnold, F. H. & Haymore, B. L. (1991) *Science* **252**, 1796–1797.
13. Sharon, N. & Lis, H. (1990) *FASEB J.* **4**, 3198–3208.
14. Derewenda, Z., Yariv, J., Helliwell, J. R., Kalb (Gilboa), A. J., Dodson, E. J., Papiz, M. Z., Wan, T. & Campbell, J. (1989) *EMBO J.* **8**, 2189–2193.
15. Bourne, Y., Roussel, A., Frey, M., Rouge, P., Fontecilla-Camps, J.-C. & Cambillau, C. (1990) *Proteins* **8**, 365–376.
16. Poretz, R. D. & Goldstein, I. J. (1968) *Arch. Biochem. Biophys.* **125**, 1034–1035.
17. Butcher, E. C. (1991) *Cell* **67**, 1033–1036.
18. Sargovi, H. U., Fitzpatrick, D., Raktabutr, A., Nakanishi, H., Kahn, M. & Greene, M. I. (1991) *Science* **253**, 792–795.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C07K 1/04, G01N 33/68, C12Q 1/68, C07K 14/00

A1

(11) International Publication Number: WO 97/22617

(43) International Publication Date: 26 June 1997 (26.06.97)

(21) International Application Number: PCT/US96/20561

(22) International Filing Date: 18 December 1996 (18.12.96)

(30) Priority Data:
08/573,786 18 December 1995 (18.12.95) US

(71) Applicant: PRAECIS PHARMACEUTICALS INCORPORATED [US/US]; One Hampshire Street, Cambridge, MA 02139-1572 (US).

(72) Inventors: BENJAMIN, Howard; 410 Marrett Road, Lexington, MA 02173 (US). FINDEIS, Mark, A.; Apartment 3A, 45 Trowbridge Street, Cambridge, MA 02138 (US). GEFTER, Malcolm, L.; 46 Baker Bridge Road, Lincoln, MA 01773 (US). MUSSO, Gary; 38 Proctor Street, Hopkinton, MA 01748 (US). SIGNER, Ethan, R.; 20 Forest Street, Cambridge, MA 02140 (US).

(74) Agents: DECONTI, Giulio, A., Jr. et al.; Lahive & Cockfield, 60 State Street, Boston, MA 02109 (US).

(81) Designated States: AU, CA, JP, European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: METHODS FOR IDENTIFYING COMPOUNDS THAT BIND TO A TARGET

(57) Abstract

Methods for identifying a compound that binds to a target are described. In general, the methods involve forming a first library comprising a multiplicity of peptides, identifying one or more peptides that bind to the target and determining a peptide motif therefrom, forming a second library comprising a multiplicity of compounds designed based on the peptide motif, selecting from the second library at least one compound that binds to the target, and determining the structure or structures of the at least one compound that binds to the target. Libraries of compounds based on a peptide motif and compounds identified by the methods of the invention are also disclosed.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AM	Armenia	**GB**	United Kingdom	**MW**	Malawi
AT	Austria	**GE**	Georgia	**MX**	Mexico
AU	Australia	**GN**	Guinea	**NE**	Niger
BB	Barbados	**GR**	Greece	**NL**	Netherlands
BE	Belgium	**HU**	Hungary	**NO**	Norway
BF	Burkina Faso	**IE**	Ireland	**NZ**	New Zealand
BG	Bulgaria	**IT**	Italy	**PL**	Poland
BJ	Benin	**JP**	Japan	**PT**	Portugal
BR	Brazil	**KE**	Kenya	**RO**	Romania
BY	Belarus	**KG**	Kyrgystan	**RU**	Russian Federation
CA	Canada	**KP**	Democratic People's Republic of Korea	**SD**	Sudan
CF	Central African Republic			**SE**	Sweden
CG	Congo	**KR**	Republic of Korea	**SG**	Singapore
CH	Switzerland	**KZ**	Kazakhstan	**SI**	Slovenia
CI	Côte d'Ivoire	**LI**	Liechtenstein	**SK**	Slovakia
CM	Cameroon	**LK**	Sri Lanka	**SN**	Senegal
CN	China	**LR**	Liberia	**SZ**	Swaziland
CS	Czechoslovakia	**LT**	Lithuania	**TD**	Chad
CZ	Czech Republic	**LU**	Luxembourg	**TG**	Togo
DE	Germany	**LV**	Latvia	**TJ**	Tajikistan
DK	Denmark	**MC**	Monaco	**TT**	Trinidad and Tobago
EE	Estonia	**MD**	Republic of Moldova	**UA**	Ukraine
ES	Spain	**MG**	Madagascar	**UG**	Uganda
FI	Finland	**ML**	Mali	**US**	United States of America
FR	France	**MN**	Mongolia	**UZ**	Uzbekistan
GA	Gabon	**MR**	Mauritania	**VN**	Viet Nam

METHODS FOR IDENTIFYING COMPOUNDS THAT BIND TO A TARGET

Background of the Invention

Recent advances in methods for producing large libraries of peptides have provided unprecedented numbers of peptides which can be screened for pharmaceutical activity. Both chemical and biological methods for synthesis of peptide libraries have been reported. For example, libraries of peptides (*e.g.*, having 10^6-10^{12} member peptides) can be displayed on the surface of bacteriophage (known as "phage display" libraries). Such peptide libraries can comprise all possible peptides of a given length (*e.g.*, every one of the twenty natural amino acid residues at each position of a hexamer), or a subset of all possible peptides. Methods for screening large libraries of peptides, to identify those peptides that bind to a target, have also been developed, such as biopanning. These screening techniques allow for the isolation from a library of one, or several, peptides that bind to a pre-selected target. By producing and screening large peptide libraries, it has become possible to rapidly search for peptides (*e.g.*, ligands) that bind to a target (*e.g.*, a receptor). Moreover, the structure of selected peptides can be determined with relative ease by standard sequencing methodologies (*e.g.*, sequencing of the peptides themselves or of a nucleic acid molecule encoding the peptide).

Despite the advantages of peptide libraries (*e.g.*, immense diversity and simple "deconvolution" of the peptide structure by sequencing), the use of this approach to identify peptides that bind a target for pharmaceutical purposes has a number of drawbacks. For example, the affinity of a selected peptide(s) for the target often is relatively low (*e.g.*, high enough to detect binding of the peptide to the target but too low for pharmaceutical potency). Moreover, peptides are not always suitable for therapeutic administration due to such problems as difficulties in formulation (due to insolubility), unfavorable pharmacokinetics and/or pharmacodynamics, and rapid degradation *in vivo*.

Alternative to peptide libraries, libraries of non-peptide chemical compounds (*e.g.*, peptidomimetics, peptide derivatives, peptide analogues, etc.) can be synthesized. Screening of a target with a non-peptide library may lead to the identification of a compound(s) with higher affinity for the target than that of a peptide selected by random peptide library screening and/or identification of a compound(s) with more desirable pharmacological properties than a peptide. However, the diversity of compounds that can be achieved by random chemical synthesis is considerably lower than that of random peptide library synthesis, thereby reducing the likelihood of identifying a high affinity target-binding compound from a randomly synthesized chemical library. An additional disadvantage of a chemical library approach to identifying molecules that bind a target is that determination of the structure of the compound(s) that binds the target (*i.e.*,

"deconvolution" of the compound structure) cannot be accomplished by a simple sequencing methodology but rather requires more complex chemical strategies, thereby limiting the number of identified compounds that can be efficiently analyzed.

Improved methods for identifying compounds that bind a target that retain the advantageous properties of both peptide library screening and chemical library screening while reducing or eliminating the disadvantageous properties of these techniques are needed.

Summary of the Invention

The present invention features methods for identifying compounds that bind a target that combine the use of peptide-based libraries with the use of chemically-based libraries such that the advantages of each approach are maintained while many of the disadvantages of using either approach alone are overcome. For example, the methods of the invention provide the diversity and ease of deconvolution of traditional peptide library screening yet also provide for the identification of compounds with high affinity for the target and desirable pharmacological properties. To optimize the benefits of both peptide-based and chemically-based libraries, the methods of the invention involve utilizing information obtained from screening a target with a first library comprising a multiplicity of peptides in the design of a second library comprising a multiplicity of chemical (*i.e.*, non-peptide) compounds. The target is then rescreened with this second library to identify compounds that bind to the target.

The methods of the invention generally involve the following steps:

a) forming a first library comprising a multiplicity of peptides;

b) selecting from the first library at least one peptide that binds to the target;

c) determining the sequence or sequences of the at least one peptide that binds to the target, thereby forming a peptide motif;

d) forming a second library comprising a multiplicity of non-peptide compounds designed based on the peptide motif;

e) selecting from the second library at least one non-peptide compound that binds to the target; and

f) determining the structure or structures of the at least one non-peptide compound that binds to the target;

thereby identifying a compound that binds to the target.

The first library is composed of peptides whose structures can be determined by standard sequencing methodologies (*e.g.*, direct sequencing of the amino acids making up the peptides or sequencing of nucleic acid molecules encoding the peptide). Thus,

the first library provides the extensive diversity of peptide libraries and the ease of deconvoluting the selected peptides. In contrast, the second, non-peptide library preferably comprises compounds that, while not peptides, are structurally related to peptides, such as peptide analogues, peptide derivatives and/or peptidomimetics. The structure of the non-peptide compounds preferably is determined by a mass spectrometric method, most preferably by tandem mass spectrometry. Since the second library is designed based on the peptide motif generated from screening the first library, many of the disadvantages of traditional chemical libraries (such as reduced diversity and more laborious deconvolution methods) are reduced or eliminated, since the second library is "biased" toward compounds that have affinity for the target. This bias in the second library for compounds having affinity for the target means that fewer compounds need to be screened as compared to a random chemically-synthesized library and, accordingly, fewer compounds need to be analyzed structurally (*i.e.*, deconvoluted).

In a preferred embodiment, compounds identified by screening of the second library have at least 10-fold higher affinity for the target than the peptides identified by screening the first library. More preferably, compounds identified by screening of the second library have at least 100-fold higher affinity for the target than the peptides identified by screening the first library. Even more preferably, compounds identified by screening of the second library have at least 1000-fold higher affinity for the target than the peptides identified by screening the first library.

The methods of the invention can further involve additional library screening steps. For example, after compounds from the second library that bind the target have been identified, a third library can be formed that comprises a multiplicity of non-peptide compounds designed based on the structure or structures of the non-peptide compounds identified from the second library. The target can be rescreened with the third library to identify additional compounds that binds to the target.

Another aspect of the invention pertains to a library comprising a multiplicity of non-peptide compounds designed based on a peptide motif, wherein the peptide motif is determined by selecting from a peptide library at least one peptide that binds to a target, determining the sequence or sequences of the at least one peptide that binds to the target and determining a peptide motif.

Yet another aspect of the invention pertains to compounds identified by a method of the invention. In a preferred embodiment, the compound is a peptidomimetic. In other preferred embodiments, the compound that binds to a target has a binding affinity for the target, expressed as an apparent K_d, EC_{50} or IC_{50}, of at least about 10^{-7} M, more preferably at least about 10^{-8} M, and even more preferably at least about 10^{-9} M.

Brief Description of the Drawing

Figure 1 is a graph depicting the ability of compounds PPI-432, PPI-652 and PPI-654 to inhibit the binding of radiolabeled FGF to an anti-FGF antibody.

Detailed Description of the Invention

The present invention pertains to methods for identifying a compound that binds to a target, as well compounds identified thereby, and libraries for use in the methods of the invention. The methods of the invention involve screening a target with at least two distinct libraries. The term "target", as used herein, is intended to include molecules or molecular complexes with which compounds (*e.g.*, peptides or non-peptide compounds) can bind or interact. Exemplary targets include ligands, receptors, hormones, cytokines, antibodies, antigens, enzymes, and the like. The target can be, for example, a purified compound or a partially purified compound or it can be associated with the surface of a cell that expresses the target.

In the methods of the invention, a target is initially screened with a peptide library to generate a peptide motif for peptides that can bind to the target. Accordingly, the methods of the invention first involve:

forming a first library comprising a multiplicity of peptides;

selecting from the first library at least one peptide that binds to the target; and

determining the sequence or sequences of the at least one peptide that binds to the target, thereby generating a peptide motif.

The term "peptides", as used herein with regard to libraries, is intended to include molecules comprised only of natural amino acid residues (*i.e.*, alanine, arginine, aspartic acid, asparagine, cysteine, glutamic acid, glutamine, glycine, histidine, isoleucine, leucine, lysine, methionine, phenylalanine, proline, serine, threonine, tryptophan, tyrosine and valine) linked by peptide bonds, or other residues whose structures can be determined by standard sequencing methodologies (*e.g.*, direct sequencing of the amino acids making up the peptides or sequencing of nucleic acid molecules encoding the peptide). The term "peptide" is not intended to include molecules structurally related to peptides, such as peptide derivatives, peptide analogues or peptidomimetics, whose structures cannot be determined by standard sequencing methodologies but rather must be determined by more complex chemical strategies, such as mass spectrometric methods.

The term "multiplicity", as used herein, refers to a plurality of different molecules (*e.g.*, peptides or non-peptide compounds). Thus a "library comprising a multiplicity of peptides" refers to a library of peptides comprising at least two different peptide members. In preferred embodiments, libraries of peptides useful in the present

invention include at least about 10^3 different peptides, more preferably at least about 10^6 different peptides and even more preferably at least about 10^9 different peptides. Depending on the length of the peptide members and the efficiency of synthesis, library diversity as high as about 10^{12} different peptides or even about 10^{15} different peptides may be achievable. A library comprising a multiplicity of peptides for use in the methods of the invention can comprise all possible peptides of a specified length (*i.e.*, a "complete" random library wherein each position of the peptide can be any one of the twenty natural amino acid residues, *e.g.*, all possible hexapeptides). Alternatively, a peptide library can include only a subset of all possible peptides of a specified length by having non-degenerate positions within the peptide library (*i.e.*, one or more positions within the peptide which are occupied by only one, or a few, different amino acid residue(s) within each peptide member of the library). Moreover, as the peptide length increases, it may not be possible to achieve every possible peptide permutation within the library. Preferably, at least about 10^5 to 10^8 permutations of all possible permutations of a randomized peptide are present within the library. The length of the peptides used in the library can vary depending upon, for example, the degree of diversity desired and the particular target to be screened. For example, in different embodiments, the peptide library is made up of peptides not longer than about 30 amino acids long, not longer than about 20 amino acids long or not longer than about 12 amino acids long. Preferably, the peptide library is comprised of peptides at least 3 amino acids long, and more preferably at least 6 amino acids long.

A library comprising a multiplicity of peptides can be formed by any one of several methods known in the art. For example, in one embodiment, a multiplicity of nucleic acid molecules encoding a multiplicity of random peptides are synthesized and the nucleic acid molecules are introduced into a vector that allows for expression of the encoded peptide library. One examples of such a library is an "external" library in which the peptide library is expressed on a surface protein of a host, such as a "phage display" library (see, *e.g.*, Smith, G.P. (1985) *Science* 228:1315-1317; Parmley, S.F. and Smith, G.P. (1988) *Gene* 73:305-318; and Cwirla, S. *et al.* (1990) *Proc. Natl. Acad. Sci. USA* 87:6378-6382). As used herein, a "phage display" library is intended to refer to a library in which a multiplicity of peptides is displayed on the surface of a bacteriophage, such as a filamentous phage, preferably by fusion to a coat protein of the phage (*e.g.*, the pIII protein or pVIII protein of filamentous phage). In phage-display methods, a multiplicity of nucleic acid molecules coding for peptides is synthesized and inserted into a phage vector to provide a recombinant vector. Suitable vectors for construction of phage display libraries include fUSE vectors, such as fUSE1, fUSE2, fUSE3 and fUSE5 (Smith and Scott (1993) *Methods Enzymol.* 217:228-257). Nucleic acid molecules can

be synthesized according to methods known in the art (see, *e.g.*, Cormack and Struhl, (1993) *Science* 262:244-248), including automated oligonucleotide synthesis . Following insertion of the nucleic acid molecules into the phage vector, the vector is introduced into a suitable host cell and the recombinant phage are expressed on the cell surface after a growth period. The recombinant phage can then be used in screening assays with a target (described further below).

Another example of a peptide library encoded by a multiplicity of nucleic acid molecules is an "internal" library, wherein the peptide members are expressed as fusions with an internal protein of a host (*i.e.*, a non-surface protein) by inserting the nucleic acid molecules encoding the peptides into a gene encoding the internal protein. The internal protein may remain intracellular or may be secreted by, or recovered from, the host. Examples of internal proteins with which peptide library members can be fused include thioredoxin, staphnuclease, lac repressor (LacI), GAL4 and antibodies. An internal library vector is preferably a plasmid vector. In one example of an internal library, referred to as a two-hybrid system (see *e.g.*, U.S. Patent No. 5,283,173 by Field; Zervos *et al.* (1993) *Cell* 72:223-232; Madura *et al.* (1993) *J. Biol. Chem.* 268:12046-12054; Bartel *et al.* (1993) *Biotechniques* 14:920-924; and Iwabuchi *et al.* (1993) *Oncogene* 8:1693-1696), nucleic acid molecules encoding a multiplicity of peptides are inserted into a plasmid encoding the DNA binding domain of GAL4 (GAL4db) such that a library of GAL4db-peptide fusion proteins are encoded by the plasmid. Yeast cells (*e.g.*, *Saccharomyces cerevisiae* YPB2 cells) are transformed simultaneously with the plasmid encoding the library of GAL4db-peptide fusion proteins and a second plasmid encoding a fusion protein composed of the target fused to the activation domain of GAL4 (GAL4ad). When the GAL4ad-target interacts with a GAL4db-peptide library member, the two domains of the GAL4 transcriptional activator protein are brought into sufficient proximity as to cause transcription of a reporter gene or a phenotypic marker gene whose expression is regulated by one or more GAL4 operators.

In another example of an internal library (see *e.g.*, U.S. Patents 5,270,181 and 5,292,646, both by McCoy), nucleic acid molecules encoding a multiplicity of peptides are inserted into a plasmid encoding thioredoxin such that a library of thioredoxin-peptide fusion proteins are encoded by the plasmid. The plasmid is introduced into a bacterial host cell where the thioredoxin-peptide fusion proteins are expressed cytoplasmically. The fusion proteins can be selectively released from the host cells (*e.g.*, by osmotic shock or freeze-thaw procedures) and recovered for use in screening assays with a target.

In yet another example of an internal library (described further in Cull, M.G. *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:1865), nucleic acid molecules encoding a

multiplicity of peptides are inserted into a gene encoding LacI to create a fusion gene encoding a fusion protein of LacI and the peptide library members. The plasmid encoding the fusion protein library members is designed such that the fusion proteins bind to the plasmid (*i.e.*, a plasmid encoding the LacI fusion proteins includes *lac* operator sequences to which LacI binds) such that the fusion proteins and the plasmids encoding them can be physically linked. Following expression of the fusion proteins in host cells, the cells are lysed to liberate the fusion protein and associated DNA, and the library is screened with an immobilized target. Fusion proteins that bind to the target are recovered and the associated DNA is reintroduced into cells for amplification and sequencing, thus allow for determination of the peptide sequence encoded by the DNA.

Alternative to forming a peptide library by synthesizing a multiplicity of nucleic acid molecules encoding the peptide library members, a multiplicity of peptides can be synthesized directly by standard chemical methods known in the art. For example, a multiplicity of peptides can be synthesized by "split synthesis" of peptides on solid supports (see, *e.g.*, Lam, K.S. *et al.* (1993) *Bioorg. Med. Chem. Lett.* 3:419-424). Other exemplary chemical syntheses of peptide libraries include the pin method (see, *e.g.*, Geysen, H.M. *et al.* (1984) *Proc. Natl. Acad. Sci. USA* 81:3998-4002); the tea-bag method (see, *e.g.*, Houghten, R.A. *et al.* (1985) *Proc. Natl. Acad. Sci. USA* 82:5131-5135); coupling of amino acid mixtures (see, *e.g.*, Tjoeng, F.S. *et al.* (1990) *Int. J. Pept. Protein Res.* 35:141-146; U.S. Patent 5,010,175 to Rutter *et al.*); and synthesis of spatial arrays of compounds (see, *e.g.*, Fodor, S.P.A. *et al.* (1991) *Science* 251:767). Peptide libraries formed by direct synthesis of the peptide library members preferably are bound to a solid support (*e.g.*, a bead or pin, wherein each bead or pin is linked to a single peptide moiety) to facilitate separation of peptides that bind a target from peptides that do not bind a target.

A particularly preferred peptide library for use in the methods of the invention is an anchor library as described in U.S. Patent Application Serial No. 08/479,660, entitled *Anchor Libraries and Identification of Peptide Binding Sequences*, and corresponding PCT Application No. PCT/US96/09383, the entire contents of both of which are expressly incorporated herein by reference. As used herein, the term "anchor library" refers to a peptide library in which the peptides have non-continuous regions of random amino acids separated by specifically designated amino acid residues. Anchor libraries are therefore subsets of a complete library of a specified length. Anchor libraries can be used to identify essential contacts between a ligand and a target, and have the advantage that only a subset of all possible peptides need be synthesized and screened. In a preferred embodiment, an anchor library is made up of peptides about 16 amino acids long. An anchor library can be prepared by genetic means (*e.g.*, by synthesizing a

multiplicity of nucleic acid molecules encoding a multiplicity of anchor peptides) or by chemical means (*e.g.*, by directly synthesizing a multiplicity of anchor peptides).

Once the peptide library has been formed, a target of interest is screened with the peptide library to identify one or more library members that bind to the target. Peptides that bind a target can be selected according to known methods, such as biopanning of an immobilized target with a phage display library. In one embodiment, a biotinylated target is immobilized on a streptavidin-coated surface either before or after contacting the target with a peptide library and unbound peptides are removed by washing. Peptide libraries bound to a solid support can be screened by, for example, contacting the peptides immobilized on the solid support with a labeled target and detecting the labeled target bound to library members or, alternatively, by releasing the peptides from the solid support and assaying the resulting solution (see, *e.g.*, Ohlmeyer, M.H.J. *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:10922:10926).

Following selection of one or more peptide library members that bind to the target, the amino acid sequence of the peptide is determined according to standard methods. For example, in one embodiment, the amino acid sequence of the peptide is determined by determining the nucleotide sequence of a nucleic acid molecule encoding the peptide and translating the encoded peptide using the genetic code. Nucleotide sequencing can be performed by standard methods (*e.g.*, dideoxynucleotide sequencing or Maxam-Gilbert sequencing, either manually or using automated nucleic acid sequencers). Alternatively, in another embodiment, the amino acid sequence of the selected peptide(s) is determined by direct amino acid sequencing of the peptide (*e.g.*, by Edman microsequencing, either manually or using automated peptide sequencers).

Once the sequence(s) of the peptide(s) that bind the target selected from the first library has been determined, a peptide motif is generated based on these sequences. As used herein, the term "peptide motif" is intended to include an amino acid consensus sequence that represents preferred amino acid residues within a peptide that are sufficient or essential for binding of the peptide to the target. Typically, the simplest way to generate a peptide motif is to compare the amino acid sequences of all peptides selected from screening a target with the first peptide library and define a peptide motif based on one or more amino acid residues that are conserved within at least two of the selected peptides. If only a single peptide is selected from the initial peptide library screening, the amino acid sequence of this peptide can constitute a peptide motif. Alternatively, when multiple peptides are selected from the initial peptide library screening, the amino acid sequences of each of the selected peptides are optimally aligned and amino acid residues conserved among two or more of the selected peptides can constitute the peptide motif. In addition to, or alternative to, direct alignment and

analysis of the primary amino acid sequence of the selected peptides. a peptide motif can be generated by more sophisticated structural analysis of the selected peptides. For example, molecular modelling programs can be employed to determine structural motifs present in the selected peptide(s). Examples of such structural motifs include α-helix, β-turns, and the like (see, *e.g.*, A. Fersht (1985) "Enzyme Structure and Mechanism", 2nd ed., W.H. Freeman and Co., New York). Computer modelling can also be used to calculate properties of active peptides such as hydrophobicity, steric bulk, stacking interactions, dipole moment, and the like. Any of the above-mentioned properties can be included when generating a peptide motif.

In the methods of the invention, after a peptide motif has been generated for a target of interest based on screening of the first library, the target is rescreened with a second, non-peptide library that is designed based on the peptide motif. The second library can be composed of compounds that are designed to have improved properties compared to the peptides selected from screening of the first library, such as increased affinity for the target (*e.g.*, predicted by computer modelling of the target with non-peptide compounds designed based on the peptide motif) and/or improved pharmacological properties, such as increased solubility, decreased susceptibility to proteolytic degradation, increased biodistribution and the like. Accordingly, the methods of the invention further comprise the steps of:

forming a second library comprising a multiplicity of non-peptide compounds designed based on the peptide motif;

selecting from the second library at least one non-peptide compound that binds to the target; and

determining the structure or structures of the at least one non-peptide compound that binds to the target.

The term "non-peptide compounds", as used herein, is intended to include compounds comprising at least one molecule other than a natural amino acid residue, wherein the structures of the compounds cannot be determined by standard sequencing methodologies but rather must be determined by more complex chemical strategies, such as mass spectrometric methods. Preferred non-peptide compounds are those that, although not composed entirely of natural amino acid residues, are nevertheless related structurally to peptides, such as peptidomimetics, peptide derivatives and peptide analogues. As used herein, a "derivative" of a compound X (*e.g.*, a peptide) refers to a form of X in which one or more reactive groups on the compound have been derivatized with a substituent group. Examples of peptide derivatives include peptides in which an amino acid side chain, the peptide backbone, or the amino- or carboxy-terminus has been derivatized (*e.g.*, peptidic compounds with methylated amide linkages). As used herein

an "analogue" of a compound X refers to a compound which retains chemical structures of X necessary for functional activity of X yet which also contains certain chemical structures which differ from X. An example of an analogue of a naturally-occurring peptide is a peptide which includes one or more non-naturally-occurring amino acids. As used herein, a "mimetic" of a compound X refers to a compound in which chemical structures of X necessary for functional activity of X have been replaced with other chemical structures which mimic the conformation of X. Examples of peptidomimetics include peptidic compounds in which the peptide backbone is substituted with one or more benzodiazepine molecules (see *e.g.*, James, G.L. *et al.* (1993) *Science* 260:1937-1942) and "retro-inverso" peptides (see U.S. Patent No. 4,522,752 by Sisto), described further below.

The term mimetic, and in particular, peptidomimetic, is intended to include isosteres. The term "isostere" as used herein is intended to include a chemical structure that can be substituted for a second chemical structure because the steric conformation of the first structure fits a binding site specific for the second structure. The term specifically includes peptide back-bone modifications (*i.e.*, amide bond mimetics) well known to those skilled in the art. Such modifications include modifications of the amide nitrogen, the α-carbon, amide carbonyl, complete replacement of the amide bond, extensions, deletions or backbone crosslinks. Several peptide backbone modifications are known, including $\psi[CH_2S]$, $\psi[CH_2NH]$, $\psi[CSNH_2]$, ψ[NHCO], $\psi[COCH_2]$, and ψ [(E) or (Z) CH=CH]. In the nomenclature used above, ψ indicates the absence of an amide bond. The structure that replaces the amide group is specified within the brackets. Other examples of isosteres include peptides substituted with one or more benzodiazepine molecules (see *e.g.*, James, G.L. *et al.* (1993) *Science* 260:1937-1942), peptoids (R.J. Simon *et al.* (1992) *Proc. Natl. Acad. Sci. USA* 89:9367-9371), and the like.

Other possible modifications of peptides include an N-alkyl (or aryl) substitution (ψ[CONR]), backbone crosslinking to construct lactams and other cyclic structures, or retro-inverso amino acid incorporation (ψ[NHCO]). By "inverso" is meant replacing L-amino acids of a sequence with D-amino acids, and by "retro-inverso" or "enantio-retro" is meant reversing the sequence of the amino acids ("retro") and replacing the L-amino acids with D-amino acids. For example, if the parent peptide is Thr-Ala-Tyr, the retro modified form is Tyr-Ala-Thr, the inverso form is thr-ala-tyr, and the retro-inverso form is tyr-ala-thr (lower case letters refer to D-amino acids). Compared to the parent peptide, a retro-inverso peptide has a reversed backbone while retaining substantially the original spatial conformation of the side chains, resulting in a retro-inverso isomer with a topology that closely resembles the parent peptide. See Goodman *et al.* "*Perspectives in*

Peptide Chemistry" pp. 283-294 (1981). See also U.S. Patent No. 4,522,752 by Sisto for further description of "retro-inverso" peptides.

Approaches to designing peptide analogues, derivatives and mimetics are known in the art. For example, see Farmer, P.S. in Drug Design (E.J. Ariens, ed.) Academic Press, New York, 1980, vol. 10, pp. 119-143; Ball, J.B. and Alewood, P.F. (1990) *J. Mol. Recognition* 3:55; Morgan, B.A. and Gainor, J.A. (1989) *Ann. Rep. Med. Chem.* 24:243; and Freidinger, R.M. (1989) *Trends Pharmacol. Sci.* 10:270.

The second, non-peptide library can be formed by methods known in the art for combinatorial synthesis of organic compounds. For example, a second library comprising compounds that include modified amino acids (for example, D-amino acids or synthetic amino acids such as phenylglycine) can be synthesized by techniques used for the synthesis of peptide libraries (*e.g.*, solid support methods described *supra*). Other organic molecules that have been synthesized on solid supports include benzodiazepines (B.A. Bunin and J.A.A. Ellman (1992) *J. Am. Chem. Soc.* 114:10997-10998), peptoids (R.N. Zuckermann *et al.* (1992) *J. Am. Chem. Soc.* 114:10646-10647), peptidyl phosphonates (D.A. Campbell and J.C. Bermak (1994) *J. Org. Chem.* 59:658-660), vinylogous polypeptides (M. Hagihara *et al.* (1992) *J. Am. Chem. Soc.* 114:6568-6570), and the like. An alternative synthetic scheme for chemical libraries involves synthesis of compounds on resin beads wherein a coding moiety corresponding to each addition in the synthesis is also coupled to the bead (see *e.g.*, Brenner, S. and Lerner, R.A. (1992) *Proc. Natl. Acad. Sci. USA* 89:5181-5183; Ohlmeyer, M.H.L. *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:10922:10926; Still *et al.*, PCT publication WO 94/08051). In a preferred embodiment, the second library comprises compounds which include at least one peptide bond (*i.e.*, amide bond). In a preferred embodiment, the second library is a library of peptidomimetics.

Preferably, the second library comprises at least about 10^2 different compounds, more preferably at least 10^4 different compounds, and still more preferably at least 10^6 different compounds. Depending upon the size of the non-peptide compounds in the library and the efficiency of synthesis, it may be possible to achieve a second library comprising as many as 10^8 different compounds or even 10^{10} different compounds.

After formation of the second library, the target of interest is screened with the second library, *e.g.*, by the screening methods described above for screening the first library. One or more non-peptide compounds that bind to the target are thereby selected. Preferably, a non-peptide compound selected from the second library that binds to a target has a binding affinity for the target, expressed as an apparent K_d (dissociation constant), EC_{50} (concentration needed for 50% effective binding) or IC_{50} (concentration needed for 50% inhibition of binding of another compound that binds to the target) of at

least about 10^{-7} M, more preferably at least about 10^{-8} M, and even more preferably at least about 10^{-9} M. In a preferred embodiment, compounds identified by screening of the second library have at least 10-fold higher affinity for the target than the peptides identified by screening the first library. More preferably, compounds identified by screening of the second library have at least 100-fold higher affinity for the target than the peptides identified by screening the first library. Even more preferably, compounds identified by screening of the second library have at least 1000-fold higher affinity for the target than the peptides identified by screening the first library.

Following selection of one or more compounds from the second library that bind to the target, the structure of the selected compound(s) is determined. In a preferred embodiment, the structure of the non-peptide compound(s) is determined by the use of a mass spectrometric method. Mass spectrometric methods allow for the rapid, inexpensive, and highly accurate identification of the structure of a compound based on the mass of the compound and on fragments of the compound generated in the mass spectrometer. A preferred mass spectrometric technique is tandem mass spectrometry, sometimes denoted "MS/MS". In tandem mass spectrometry, a sample compound is first ionized and the molecular ion determined. The molecular ion is then cleaved into several smaller fragments, which are then mass-analyzed. The use of mass spectrometry to identify the structure of high-molecular weight compounds, including peptides, has been reported (see, *e.g.*, R.S. Youngquist *et al.* (1995) *J. Am. Chem. Soc.* 117:3900; B.J. Egner *et al.* (1995) *J. Org. Chem.* 60:2652-2653). It is believed that tandem mass spectrometry is especially useful for the analysis of non-peptide compounds that contain one or more peptide bonds (*e.g.*, peptide derivatives, peptide analogues and/or peptidomimetics) because the peptide bond can be cleaved in the spectrometer to produce fragments that can be analyzed to identify particular subunits of the compound. In certain alternative embodiments, it may be possible to analyze at least a portion of a non-peptide compound by direct amino acid sequencing, *e.g.*, by Edman degradation (*e.g.*, where the non-peptide compound comprises a peptide portion). Alternatively, in embodiments in which the second library is synthesized in an array (*e.g.*, on pins or in an array on a solid surface, *e.g.*, a "chip"), the structure of the compound can be determined by the position the compound occupies in the array. In yet other embodiments, in which the second library is an encoded library (*i.e.*, a library in which the structure of the chemical compound has been encoded on a bead, as described in Brenner, S. and Lerner, R.A. (1992) *Proc. Natl. Acad. Sci. USA* 89:5181-5183; Ohlmeyer, M.H.L. *et al.* (1993) *Proc. Natl. Acad. Sci. USA* 90:10922:10926; and Still *et al.*, PCT publication WO 94/08051), the structure of the compound can be determined by decoding the encoding moiety.

In a particularly preferred embodiment of the method of the invention, the first (peptide) library is a phage display library, and the non-peptide compound(s) of the second library that bind to the target are analyzed by tandem mass spectrometry. In another particularly preferred embodiment of the methods of the invention, the first (peptide) library is an anchor library, and the compound(s) of the second library that bind to the target are analyzed by tandem mass spectrometry.

The skilled artisan will appreciate that the compound or compounds identified from the second library can be used as a basis for forming further libraries that can be used for further screening of the target. That is, the information gained from the screening of the second library can be used to design another motif, for example a modified peptide motif (*e.g.*, a motif based on the structure of peptide derivatives, peptide analogues and/or peptidomimetics), and a subsequent, third library can be formed comprising compounds designed based on the motif generated from the screening of the second library. The target is then screened with the third library and active compounds identified as previously described herein. This process can be repeated until a compound with a desired binding affinity for the target is obtained.

Another aspect of the invention pertains to a compound identified by the method of the invention. In preferred embodiments, the compound is a peptidomimetic, peptide derivative or peptide analogue. Preferably, a compound identified by the method of the invention has a binding affinity for the target, expressed an an apparent K_d (dissociation constant), EC_{50} (concentration needed for 50% effective binding) or IC_{50} (concentration needed for 50% inhibition of binding of another compound that binds to the target) of at least about 10^{-7} M, more preferably at least about 10^{-8} M, and even more preferably at least about 10^{-9} M. The binding affinity of a compound for a particular target can be determined by standard methods for determining K_ds, EC_{50}s or IC_{50}s.

Another aspect of the invention pertains to a library comprising a multiplicity of non-peptide compounds designed based on a peptide motif, wherein the peptide motif is determined by selecting from a peptide library at least one peptide that binds to a target, determining the sequence or sequences of at least one peptide, preferably multiple peptides, that binds to the target and determining a peptide motif. A library of non-peptide compounds based on a peptide motif can be synthesized by the methods previously described herein. In a preferred embodiment, the non-peptide compounds of the library are peptidomimetics. Additionally or alternatively, the non-peptide compounds can be peptide derivatives and/or peptide analogues. Preferably, the library comprises at least about 10^2 compounds, more preferably at least about 10^4 compounds and even more preferably at least about 10^6 compounds. In one embodiment, the

multiplicity of non-peptide compounds are attached to a solid support, such as a plurality of resin beads.

This invention is further illustrated by the following examples which should not be construed as limiting. The contents of all references, patents and published patent applications cited throughout this application are hereby incorporated by reference.

EXAMPLE 1

In this example, the method of the invention is used to identify one or more compounds that bind to a target that is expressed on the surface of a cell, the luteinizing hormone releasing hormone receptor (LHRH-R), a member of the G-protein coupled, seven transmembrane receptor superfamily.

Construction of the First Library

A phage anchor library comprising a multiplicity of peptides is used as the first library in the method. The anchor library is comprised of peptides having random amino acid residues distributed throughout domains of alanine (Ala) and/or glycine (Gly) residues. For example, the anchor library can be composed of peptides that are sixteen amino acid residues in length and have the amino acid sequence:

$$X^1(Ala/Gly)_4X^2(Ala/Gly)_4X^3(Ala/Gly)_4X^4$$

wherein X^1, X^2, $X^3$and X^4 can be any amino acid residue and each can be the same or different from the others.

To prepare the anchor library, a multiplicity of oligonucleotides encoding the peptides are synthesized by standard methods, such as the split synthesis method (See *e.g.*, Cormack and Struhl (1993) *Science* 262:244-248). Synthesis of oligonucleotides for construction of anchor libraries also is described further in U.S. Patent Application Serial No.08/479,660, entitled *Anchor Libraries and Identification of Peptide Binding Sequences*, and corresponding PCT Application No. PCT/US96/09383, the entire contents of both of which are expressly incorporated herein by reference.

Following synthesis, assembled oligonucleotide inserts are cloned into the pfUSE5 phage vector (Smith and Scott (1993) *Methods in Enzymology* 217:228-257), which allows for expression of the encoded peptides as fusions with the pIII phage coat protein. The vector (30 μg) is prepared by cleaving with 200 units of endonuclease SfiI in 500 μl of restriction buffer (Buffer #2 from New England BioLabs (NEB), Beverly, MA) for 10 hours. The reaction is terminated with addition of 15 mM EDTA, followed by phenol/chloroform extraction. The vector DNA is recovered by isopropanol precipitation, resuspended in 500 μl of Tris-EDTA (TE) buffer and recovered by ethanol

precipitation. The phage vector is ligated to the assembled oligonucleotide inserts at 5 μg/ml vector and three-fold excess assembled insert in ligation buffer (NEB) with 100 units of T4 DNA ligase at 10° C for 16 hours. DNA is purified from the ligation buffer by phenol/chloroform extractions, followed by ethanol precipitations and resuspension in TE buffer.

DNA from the ligation reaction is transformed into electrocompetent MC1061 bacterial host cells (Wertman *et al.* (1986) *Gene* 49:253-262) using 0.5 μg of DNA per 100 μl of cells using 0.2 cm electroporator cells and a BioRad electroporator set at 25 μF, 2.5 KV and 200 ohms. Shocked cells are recovered in SOC media, grown out at 37 °C for 20 minutes and inoculated into LB broth containing 20 μg/ml tetracycline.

Library phage released from the transformed bacterial host cells are isolated after growing the bacterial cells for 16 hours. Phage are separated from cells by centrifugation at 4 °C at 4.2 K rpm for 30 minutes in a Beckman J6 centrifuge, followed by a second centrifugation of the supernatant at 4.2K rpm for 30 minutes. Phage are precipitated with the addition of 150 ml of 16.7 % polyethyleneglycol (PEG)/3.3 M NaCl per liter of supernatant. Mixed solutions are incubated at 4 °C for 16 hours. Precipitated phage are collected by centrifugation at 4.2K rpm in a J6 centrifuge, followed by resuspension in 40 ml of Tris-buffered saline (TBS). Resuspended phage are precipitated again with the addition of 4.5 ml of PEG solution for 4 hours. Phage are collected at 5K rpm in a Beckman JA20 centrifuge at 4° C. Phage are suspended in 7 ml of TBS and brought to 1.3 mg/ml density by the addition of 1 gm of CsCl per 2.226 gm of aqueous solution. Phage are subjected to equilibrium centrifugation in a type 80 rotor at 45K rpm for 40 hours. Phage bands are isolated, diluted 20-fold with TBS and pelleted at 40K rpm in a type 50 rotor. Pellets are resuspended in 0.7 ml of TBS and as is in screening assays, described below, at approximately 3×10^{13} phage/ml.

Screening of the First Library-

To identify members of the phage anchor library that bind to LHRH-R, monolayers of cells expressing LHRH-R (such as CHO, COS or SF9 cells transfected to express LHRH-R) adhered to culture dishes are biopanned with the phage library. The phage (in TBS) are incubated with the cells for 1 hour at 4 °C and non-specific phage are removed by washing the cell monolayer with PBS containing 2 % milk or 1 % BSA or 10 % serum for a total of 7 washes over 30 minutes. The remaining phage that are bound to the cells (by way of binding to LHRH-R on the surface of the cells) are recovered by elution with 100 μM glycine, pH 2.2 for 10 minutes. Eluted phage are neutralized with 1 M Tris base.

Eluted phage are amplified by infection of log phase K91 *E. coli* (Lyons and Zinder (1972) *Virology* 49:45-60; Smith and Scott (1993) *Methods in Enzymology* 217:228-257). Approximately 10^5 phage are amplified by infecting an equal volume of K91 cells with phage at 22 °C for 10 minutes. Infected cells are diluted into 1 ml of LB broth for 30 minutes at 37 °C, followed by an additional dilution with 9 ml of LB containing 20 μg/ml tetracycline and grown overnight. Phage are then separated from cells by centrifugation and purified by PEG precipitation and resuspended at 10^{12} phage/ml.

To further enrich for peptides that specifically bind to LHRH-R, amplified phage can be subjected to two additional rounds of biopanning using different cell types expressing LHRH-R in each round of panning and using the binding and amplification conditions described above.

Generation of a Peptide Motif

After biopanning, individual phage are isolated and sequenced to reveal the DNA sequence that encodes for the displayed peptide in each selected phage. Sequencing is performed by standard methods (*e.g.*, dideoxy sequencing using Sequenase 2.0, United States Biochemical Co., Cleveland OH, according to the manufacturer's protocol).

After obtaining the DNA sequences encoding the selected peptides, the DNA sequences are optimally aligned to generate a peptide motif. The peptide motif is determined from the amino acid residues that are conserved in at least two of the selected peptides. For example, if biopanning of the anchor library leads to selection of four peptides having the following amino acid sequences (standard three-letter abbreviations are used for amino acids):

Ser-$(Ala/Gly)_4$-Arg-$(Ala/Gly)_4$-Leu-$(Ala/Gly)_4$-Met (SEQ ID NO: 1)
Ser-$(Ala/Gly)_4$-Lys-$(Ala/Gly)_4$-Leu-$(Ala/Gly)_4$-Gln (SEQ ID NO: 2)
Phe-$(Ala/Gly)_4$-Arg-$(Ala/Gly)_4$-Leu-$(Ala/Gly)_4$-Thr (SEQ ID NO: 3)
Ser-$(Ala/Gly)_4$-Asn-$(Ala/Gly)_4$-Leu-$(Ala/Gly)_4$-Ile (SEQ ID NO: 4)

a peptide motif can be generated having the amino acid sequence:

Ser-$(Ala/Gly)_4$-Arg-$(Ala/Gly)_4$-Leu-$(Ala/Gly)_4$-Xaa (SEQ ID NO: 5)

(wherein Xaa can be any amino acid residue).

Construction of a Second Library

Based on the peptide motif generated from screening the target with the first library, a second library comprising a multiplicity of non-peptide compounds is synthesized by standard chemical synthesis methods (see *e.g.*, Youngquist, R.S. *et al.* (1995) *J. Am. Chem. Soc.* 117:3900-3906; Till, J.H. *et al.* (1994) *J. Biol. Chem.* 269:7423-7428; Berman, J. *et al.* (1992) *J. Biol. Chem.* 267:1434-1437). The non-peptide compounds of the library are designed to mimic the peptide motif. For example, to create a non-peptide library based on the peptide motif described above, amino acid derivatives, analogues or mimetics of Ser at position 1, Arg at position 6, Leu at position 11 and/or Xaa at position 16 can be incorporated into the library. Derivatives, analogues and/or mimetics of the repeating Ala/Gly structure can also be incorporated into the library.

One example of a second library synthesized based on the above-described peptide motif is an analog library in which the serine at position 1 of the motif is substituted with homoserine, cyanoalanine, isoglutamine or isoasparagine, the arginine at position 6 of the motif is substituted with citrulline, isopropyllysine, homoarginine, ornithine, homocitrulline, diaminoproprionic acid, aminobenzoic acid or nitroarginine, the leucine at position 11 of the motif is substituted with NorLeu, BuGlycine, cyclohexylalanine, norval, aminobutyrl or various N-methyl aliphatic amino acids and the Xaa at position 16 of the motif is combinatorially derived from the twenty natural amino acids or standard analogs thereof.

Another example of a second library synthesized based on the above-described peptide motif is a library constructed to probe the stereochemical specificity of compounds that bind to the target by alternating D- and L-amino acids in the library. In this case, the library is constructed using the following L-amino acids: Glu, Arg, Asn, Thr, Val, Pro, Met, Tyr and His; and the following D-amino acids: Asp, Lys, Gln, Ser, Cha, Ala, Phe and Trp. The library also contains glycine. This library can define the role of D or L stereochemistry within the selected peptide motif.

Yet another example of a second library synthesized based on the above-described peptide motif is a mimetic library, wherein reduced amide mimetics are incorporated into the compounds of the library via the use of appropriate amino acid aldehyde precursors and the solid phase reductive amination procedure for assembly (Sasaki and Coy (1987) *Peptides* 8:119-120). Mimetics can be incorporated at one site or multiple sites within the library. Appropriate positions include sites within a peptide motif containing an aliphatic or aromatic residue, such as the leucine at position 11 of the above-described peptide motif.

Once synthesized, the library is dissolved in 1-5 % dimethylsulfoxide (DMSO) in water and used in screening assays as described below.

Screening of the Second Library

To identify members of the second library that bind to LHRH-R, membranes of CHO cells that have been transfected to express LHRH-R on their surface are prepared. One liter quantities of CHO-LHRH-R cells (*e.g.*, 10^9 cells/liter) are grown and harvested. The cells are lysed with a nitrogen bomb (see *e.g.*, Autuori, F. *et al.* (1982) *J. Cell Sci.* 57:1-13). 25 ml of a washed cell suspension (in an isomolar Hanks balanced salt solution/20 mM HEPES buffer, pH 7.4) is placed in a nitrogen bomb at 4 °C with continuous stirring by a magnetic stir bar, and the pressure is adjusted to 4-500 psi, followed by continuous stirring for 20 minutes. Pressure is released into a 50 ml plastic centrifuge tube containing a 100X cocktail of protease inhibitors (0.5 mM PMSF, 10 µg/ml benzamidine, 1 µg/ml leupeptin, final concentrations). The homogenate is centrifuged for 1 hour at 5.000 Xg. The supernatant is subjected to ultracentrifugation at 50.000 Xg for one hour. The final pellet is resuspended to a concentration of about 1 mg/ml and aliquots are frozen in liquid nitrogen until used in binding assays, at which time the aliquots are thawed.

Binding reactions are set up in 12 x 75 mm polypropylene test tubes containing a sample of the second library (final concentration of 10 mg/ml), binding buffer (final concentrations: 10 mM Tris, 0.05 % bovine serum albumin, pH 7.4) and a sample of the cell membrane preparation (approximately 10^7 cell equivalents per tube) in a total volume of 500 µl. The binding reaction is incubated on ice for 90 minutes. The binding reaction is terminated by fast filtration binding of the mixture using a 12-well cell harvester (Millipore, Milford, MA). Filters (Whatman glass-fiber filters GF/C) are prewashed three times with 300 µl of 10 mM HEPES, 0.01 % sodium azide. 3 ml of HEPES buffer is added to each binding reaction tube and the contents of the tube are poured over the filter in the fast filtration binding apparatus. Two additional aliquots of buffer are added to each tube and poured over the filter. Compounds from the second library that bind to the LHRH-R membrane preparation are retained on the filter, whereas compound that do not bind to the LHRH-R membrane preparation are removed.

Identification of Compounds that Bind the Target

Compounds from the second library that bind to the LHRH-R membrane preparation are recovered from the filter of the fast filtration binding apparatus. The structures of the selected compounds are determined by tandem mass spectrometry (see *e.g.*, Hunt, D.F., *et al.* (1985) *Anal. Chem.* 57:765-768; Hunt, D.F. *et al.* (1986) *Proc.*

Natl. Acad. Sci. USA 83:6233-6237; Hunt, D.F. *et al.* (1987) *Proc. Natl. Acad. Sci. USA* 84:620-623; Biemann, K. (1990) *Methods in Enzymology* 193:455-479; Arnott, D. *et al.* (1993) *Clin. Chem.* 39:2005-2010; Metzger, J.W. *et al.* (1994) *Anal. Biochem.* 219:261-277; Brummel, C.L. *et al.* (1994) *Science* 264:399-402).

EXAMPLE 2

In this example, the method of the invention was used to identify compounds that bind to a fibroblast growth factor (FGF) binding protein, namely an anti-FGF monoclonal antibody. Starting with biologically generated peptide libraries in bacteriophage M13, important amino acids for target binding were defined. Based on the information generated from the bacteriophage library, a combinatorial chemical library with a complexity of approximately 160,000 compounds was designed and synthesized. This combinatorial library contained biased amino acid residues at key positions. A combination of natural amino acids (*i.e.*, the 20 amino acids encoded by DNA) and synthetic amino acids (those containing unnatural R-groups, or the D-enantiomer of a natural amino acid) were utilized in the chemical library. Non-peptide compounds that interacted with the target were recovered and were analyzed. Several of these selected non-peptide compounds were synthesized and were shown to be 10-100 fold more potent than the starting natural peptide for target binding. These non-peptide compounds contained unnatural amino acids at 3 or more positions. This example demonstrates that novel, high potency non-peptide compounds containing unnatural amino acids can be discovered using phage display library screening coupled with combinatorial chemistry library screening.

Derivation of a Peptide Sequence Motif from a Phage Display Library

A monoclonal antibody raised against human basic FGF (bFGF) was used as a biopanning target using a 7-mer peptide library in a bacteriophage M13 vector. To prepare the biopanning plates, four wells of an Immulon4 microtiter plate (Dynatech) were coated with stepavidin (1 μg/well) in 100 mM $NaHCO_3$, pH 9.5, for 2 hours at room temperature or overnight at 4°C. The strepavidin-coated wells were washed three times with phosphate-buffered saline (PBS). To each well containing strepavidin, biotinylated rat anti mouse antibody κ light chain (PharMingen) was added, at 1 μg/well, in PBS for 30 minutes to 1 hour at room temperature. The wells were then washed again with PBS three times. Non-specific binding sites were blocked with 300 μl/well PBS, 1% dry milk for 1 hour at room temperature. The wells were then washed with PBS, 0.1% dry milk three times, loaded with 50 μl/well PBS, 0.1% dry milk and stored at 4°C.

Monoclonal antibody to bFGF (Sigma) was added to four tubes in 100 µl of PBS, 0.1% dry milk at the following final concentrations: 100 nM, 25 nM, 5 nM and no antibody (control). To each tube containing anti-bFGF antibody, approximately $1x10^{10}$ phage from a seven-mer peptide library were added. The tubes were left at room temperature for 2 hours to allow for specific interaction between the phage and the anti-bFGF antibody.

After the 2 hour room temperature incubation of the phage with the antibody, the PBS, 0.1% milk buffer was removed from the wells of the strepavidin/biotinylated α-mouse complex plates and the corresponding 100 µl of each tube containing the antibody/phage mixture was added to each of the four wells . The solutions were allowed to sit in the wells at room temperature for twenty minutes. To remove unbound or non-specifically bound phage, each well was washed with cold PBS, 0.1% Tween six times, followed by cold PBS washes (six times). Specifically-bound phage were eluted with 100 µl Glycine pH 2.2 for 10 minutes at room temperature. The glycine solution was then removed from wells and added to polypropylene tubes with 6 µl 2M Tris base to neutralize. The phage eluant was titered and the fractional yield determined.

For phage amplification, phage were mixed with concentrated mid log phase *E. coli* strain K91 at room temperature for 5 to 10 minutes. One milliliter of LB was added and the cell were grown at 37°C for 30 min. Nine milliliters of LB^{tet} was then added and the cells were allowed to grow overnight at 37°C. The following morning, the bacteria were pelleted by centrifugation at 5000 rpm for 15 minutes. The supernatant was drained into a fresh 50 ml conical tube and 1.5 ml of PEG/NaCl was added. The tube was chilled at 4°C for 4 hours. The phage were pelleted by centrifugation at 8000 rpm for 30 minutes. The supernatant was drained off and the phage pellet was resuspended in 1 ml PBS. The phage amplification was titered and $1x10^{10}$ phage were used for subsequent rounds of screening as described above.

A number of related phage were selected by this biopanning procedure. The selected peptides, as deduced from DNA sequence analysis of the phage, indicated that the target protein binds to peptides containing the consensus sequence: P-x-G-H-x-K-x (SEQ ID NO: 6). Analysis of the bFGF sequence indicated that the natural epitope for the antibody is P-P-G-H-F-K-D (SEQ ID NO: 7), based on the strong similarity to the peptides selected from the phage display library. Two peptides containing this sequence were synthesized: PPI-416 = G-A-F-P-P-G-H-F-K-D-P-D-R-L (SEQ ID NO: 8) and PPI-432 = P-P-G-H-F-K-D (SEQ ID NO: 7) and tested for their ability to bind to the target. Competitive binding experiments demonstrated that PPI-416 blocked phage containing the sequence P-R-G-H-W-K-Q (SEQ ID NO: 9) from binding to the antibody. Furthermore both PPI-416 and PPI-432 blocked the binding of bFGF to the

antibody. Therefore, peptides deduced from phage biopanning bound to the target protein and blocked the interaction of bFGF. The peptide sequence information obtained from phage biopanning was applied to the synthesis of a biased combinatorial peptide library.

Design and Synthesis of Combinatorial Chemical Libraries

Sequence information derived from phage binding to the target protein, together with the known sequence of bFGF, was used in the design of a secondary biased chemical combinatorial library. Two libraries were synthesized containing a different number of fixed and variable amino acids:

	Sequence	Complexity
Library I	P-P-G-H-x-x-x (SEQ ID NO: 10)	1,600
Library II	P-P-x-x-x-x-x (SEQ ID NO: 11)	160,000

In these libraries, the first 2 or 4 amino acids were fixed to match the amino acids presumed to be involved in target interaction, and the remaining positions were variable. Natural amino acids, L-amino acids with unnatural R-groups, and D-enantiomers of natural amino acids were incorporated into the library. Accordingly, the members of these libraries are referred to herein as "non-peptide compounds". The natural and unnatural amino acids chosen at the variable positions were biased based on the sequences obtained from phage display. For example, many different amino acids were found in the last residue (carboxy-terminal) from the panned phage (position 7). Therefore, 24 different natural and unnatural amino acids were used at this position. Likewise, at the penultimate position (position 6), most phage contained a lysine residue. Therefore, this position was biased in the library synthesis by using predominantly basic natural and unnatural amino acids.

The following abbreviations for unnatural amino acids are used herein:

Abbreviation	Residue
Abu	2-amino butyric acid
Nor-Val	norvaline
Hyp	hydroxyproline
Cit	citrulline
Nal	3-(2-naphthyl)alanine
Orn	ornithine

Pal	3-(3'-pyridyl)alanine
Cha	cyclohexylalanine
Tic	1,2.3.4-tetrahydroisoquinoline-3-carboxylic acid
pF-Phe	para-fluorophenylalanine

The composition of the combinatorial library is summarized below in Table 1.

Table 1

Amino Acid	Residue MW	Library AA7	Library AA6	Library AA5	Library AA4	Library AA3	Library AA2	Library AA1
Gly	57.1	X				X		
L-Abu	85.1	X		X				
L-Pro	97.1	X				X	X	X
L-NorVal	99.1	X		X				
L-Thr	101.1	X			X			
L-Hyp	113.2	X			X	X		
L-Asp	115.1	X			X			
L-Lys	128.2	X	X		X			
L-Met	131.2	X		X				
L-His	137.2	X	X	X	X	X		
L-Phe	147.2	X		X				
L-Arg	156.2	X	X		X	X		
L-Cit	157.2	X	X		X	X		
L-Tyr	163.2	X		X				
L-Nal	197.3	X		X				
D-Ala	71.1	X		X				
D-Ser	87.1	X			X			
D-Orn	114.2	X	X		X			
D-Glu	129.1	X			X			
D-Pal	148.2	X	X	X	X	X		
D-Cha	153.2	X		X				
D-Tic	159.2	X		X		X		
D-pF-Phe	165.2	X		X				
D-Trp	186.2	X		X				

X- denotes this amino acid was used in the library construction

An additional consideration in the design of the library was to use amino acids of different molecular weight at any one position. The sequence of compounds selected from the library ultimately is determined using fragmentation and MS analysis. It is therefore important to choose natural and unnatural amino acids with distinct molecular weights to avoid ambiguity in the sequence determination.

The library was synthesized using equimolar mixtures of each natural or unnatural amino acid at each synthetic step using standard Fmoc chemistry. As a representative example, for Library II (P-P-x-x-x-x-x; SEQ ID NO: 11), the strategy involved use of 1.7 mmoles of a PEG-PAL resin (4.5 g at 0.38 mmol/g), C-terminally modified with an amide, at the start of synthesis. Synthesis using the natural and unnatural amino acids indicated in Table 1 is carried out through five cycles (*i.e.*, amino acid positions 7-3 as indicated in Table 1). At this point, 0.1 mmole of the resin is removed (1/17 of total) and the synthesis is completed with Pro for position 2 (AA2 as indicated in Table 1), Pro for position 1 (AA1 as indicated in Table 1) and an N-terminal acetylation to prepare a test library of approximately 2×10^5 complexity. For coupling, depending on the diversity of each step, individual amino acids are weighed to provide a total of 3.4 mmoles per cycle. The combined amino acids are dissolved in 20 ml of NMP so that 10 ml (equivalent) is used for the first coupling and 10 ml for the second coupling. The first coupling uses HBTU/HOBT chemistry and the second coupling uses DCC/HOAt coupling strategy. Limiting amounts of Fmoc amino acids are used at each cycle to achieve uniform distribution of each natural or unnatural amino acid. Coupling efficiency is checked with ninhydrin reagent after each cycle of synthesis to assure the reactions went to completion. Following synthesis, the library is cleaved from the resin, precipitated and lyophilized. To remove scavengers, a chromatography step is performed following lyophilization. The final library is evaluated by HPLC to assure the complexity of the library.

Library Characterization

Each complete library, containing a mixture of approximately 1600 non-peptide compounds (Library I) or 160,000 non-peptide compounds (Library II), was analyzed by MS. The mass spectrum of each library was consistent with the distribution of molecular weights predicted from the diverse amino acids used for the synthesis.

The activities of the libraries were tested using an FGF binding assay. In comparison to the peptide PPI-432 (P-P-G-H-F-K-D; SEQ ID NO: 7), the unselected Library I was approximately 10-fold less potent in blocking the binding of FGF to the

target protein. Similarly, the unselected Library II was about 1000-fold less potent that PPI-432. These data are consistent with the libraries containing a mixture of diverse non-peptide compounds whose average potency is decreased relative to the original peptide upon which the library was designed. Since Libraries I and II are highly diverse, the active non-peptide compounds should have a spectrum of potencies which cumulatively yield the average potency seen in the library. Selective enrichment of non-peptide compounds that interact with the target should be proportional to their binding affinities, allowing recovery of the most active components in the library.

Selection of Non-Peptide Compounds for Functional Analysis

The target FGF binding protein (α-FGF-mAb) was biotinylated using activated N-hydroxy-succinimide ester. The non-peptide compound library (500 μg) was incubated with the biotinylated target antibody (17.5 μg) in 500 μl phosphate buffered saline (PBS) at 4°C for 2 hours. Bound non-peptide compounds were recovered in complex with the antibody by capture on magnetic streptavidin beads as follows.

Magnetic streptavidin beads (100 μl beads per reaction) were immobilized on a magnet and washed twice with PBS. The beads were then resuspended in 100 μl PBS per sample. Washed beads (100 μl) were added to each antibody-library binding reaction and the mixture was incubated for 15 minutes at 4° C with constant rotation. The tubes were spun briefly to clear beads off the wall of the tube and the beads were then washed three times with PBS. The beads were then pelleted by centrifugation or by use of a magnet and the PBS removed. To elute bound non-peptide compounds, 100 μl of 10% acetic acid was added and the mixture was incubated at 4° C for 15 minutes. The beads were immobilized on the magnet and the supernatant, containing the eluted non-peptide compounds, was recovered.

The eluted non-peptide compounds were brought to dryness under vacuum. Recovered non-peptide compounds were dissolved in 50 μl water and tested for their ability to block the interaction of bFGF with the target protein. (Assuming 100% recovery of non-peptide compounds and 100% binding to bivalent antibody, the amount of recovered non-peptide compounds was calculated to be 200 ng). Binding experiments demonstrated that relative to the unselected library, the potency of the selected population of non-peptide compounds was increased by approximately 1000-fold.

Selection of Non-Peptide Compounds for MS and Sequence Analysis

Selection of peptide for liquid chromatography/mass spectrometry (LC/MS) analysis and sequence analysis was performed as described above, except 50 μg of

antibody and 1 mg of library was used. Selected non-peptide compounds were eluted with acetic acid and dried by lyophilization.

The selected non-peptide compounds were first analyzed by coupled LC/MS. The results showed the presence of non-peptide compounds with various chromatographic elution times. Individual fractions across the chromatographic gradient were next analyzed by MS. Different fractions contained a relatively small number of non-peptide compounds with distinct molecular masses. For example, chromatographic fraction 205-215 contained a major M/Z ion of 485.7. This doubly charged ion corresponds to a non-peptide compound with a molecular mass of approximately 971. This peak was fragmented and analyzed by tandem MS in order to deduce the structure of the non-peptide compound. The fragmentation pattern of this peptide together with the known natural and non-natural amino acids used in the library synthesis allowed for the unambiguous determination of the structure of the non-peptide compound, which structure is as follows: Pro-Pro-Gly-His-Nal-Lys-Nal (SEQ ID NO: 12). In a similar manner, the structures of several other major MS peaks were determined and are summarized below in Table 2:

Table 2: Structure of Compounds that Bind to Anti-FGF Monoclonal Antibody

	1	2	3	4	5	6	7	SEQ ID NO:
(PPI-432)	Pro	Pro	Gly	His	Phe	Lys	Asp	7
(PPI-652)	Pro	Pro	Gly	His	Nal	Lys	Nal	12
(PPI-654)	Pro	Pro	Gly	His	Nal	Lys	D-pF-Phe	13
	Pro	Pro	Gly	His	Phe	Lys	Nal	14
	Pro	Pro	Gly	His	Nal	Lys	Abu	15
	Pro	Pro	Gly	His	Nal	Lys	D-Ala	16
	Pro	Pro	Pal	x	x	x	x	17

Activity of Non-Peptide Compounds Selected from the Secondary Combinatorial Chemical Library

Several of the non-peptide compounds selected from the combinatorial library were made synthetically and compared to the starting peptide (PPI-432) for their ability to bind to the target protein. The compounds were synthesized with an amino-terminal acetyl group and a carboxy-terminal amide. The potency of two non-peptide compounds (PPI-652 and PPI-654, shown in Table 2), containing modified amino acids at positions 5 and 7, were compared to the starting peptide PPI-432 in their ability to inhibit radiolabeled FGF binding to biotinylated anti-FGF antibody. The results of this

experiment are shown in Figure 1, which demonstrate that PPI-652 and PPI-654 exhibit 10-100-fold higher binding affinity for FGF than the starting peptide PPI-432. Thus, the recovery of non-peptide compounds from diverse combinatorial libraries following selection against target correlates with the potency of those non-peptide compounds in target binding.

EQUIVALENTS

Those skilled in the art will recognize, or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described herein. Such equivalents are intended to be encompassed by the following claims.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: PRAECIS PHARMACEUTICALS INCORPORATED
(B) STREET: ONE HAMPSHIRE STREET
(C) CITY: CAMBRIDGE
(D) STATE: MASSACHUSETTS
(E) COUNTRY: USA
(F) POSTAL CODE (ZIP): 02139-1572

(ii) TITLE OF INVENTION: Methods for Identifying Compounds that Bind to a Target

(iii) NUMBER OF SEQUENCES: 17

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: LAHIVE & COCKFIELD
(B) STREET: 60 State Street, suite 510
(C) CITY: Boston
(D) STATE: Massachusetts
(E) COUNTRY: USA
(F) ZIP: 02109-1875

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: US
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: US 08/573,786
(B) FILING DATE: 18-DEC-1995

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: DeConti. Giulio A., Jr.
(B) REGISTRATION NUMBER: 31,503
(C) REFERENCE/DOCKET NUMBER: PPI-012CPPC

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: (617)227-7400
(B) TELEFAX: (617)227-5941

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2-5,7-10,12-15
(D) OTHER INFORMATION: /note= Xaa is Ala or Gly

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Ser Xaa Xaa Xaa Xaa Arg Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Met
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2-5,7-10,12-15
(D) OTHER INFORMATION: /note= Xaa is Ala or Gly

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Ser Xaa Xaa Xaa Xaa Lys Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Gln
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2-5,7-10,12-15
(D) OTHER INFORMATION: /note= Xaa is Ala or Gly

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Phe Xaa Xaa Xaa Xaa Arg Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Thr
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2-5,7-10,12-15
(D) OTHER INFORMATION: /note= Xaa is Ala or Gly

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Ser Xaa Xaa Xaa Xaa Asn Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Ile
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 16 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2-5,7-10,12-15
(D) OTHER INFORMATION: /note= Xaa is Ala or Gly

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 16
(D) OTHER INFORMATION: /note= Xaa is any amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Ser Xaa Xaa Xaa Xaa Arg Xaa Xaa Xaa Xaa Leu Xaa Xaa Xaa Xaa Xaa
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 2, 5, 7
(D) OTHER INFORMATION: /note= Xaa is any amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Pro Xaa Gly His Xaa Lys Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

Pro Pro Gly His Phe Lys Asp
1 5

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 14 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

Gly Ala Phe Pro Pro Gly His Phe Lys Asp Pro Asp Arg Leu
1 5 10

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Pro Arg Gly His Trp Lys Gln
1 5

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5-7
(D) OTHER INFORMATION: /note= Xaa is any amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Pro Pro Gly His Xaa Xaa Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3-7
(D) OTHER INFORMATION: /note= Xaa is any amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Pro Pro Xaa Xaa Xaa Xaa Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5, 7
(D) OTHER INFORMATION: /note= Xaa is Nal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Pro Pro Gly His Xaa Lys Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= Xaa is Nal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= Xaa is D-pF-Phe

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

Pro Pro Gly His Xaa Lys Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= Xaa is Nal

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

Pro Pro Gly His Phe Lys Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= Xaa is Nal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= Xaa is Abu

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

Pro Pro Gly His Xaa Lys Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:

(A) NAME/KEY: Modified-site
(B) LOCATION: 5
(D) OTHER INFORMATION: /note= Xaa is Nal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 7
(D) OTHER INFORMATION: /note= Xaa is D-Ala

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

Pro Pro Gly His Xaa Lys Xaa
1 5

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 7 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(v) FRAGMENT TYPE: internal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 3
(D) OTHER INFORMATION: /note= Xaa is Pal

(ix) FEATURE:
(A) NAME/KEY: Modified-site
(B) LOCATION: 4-7
(D) OTHER INFORMATION: /note= Xaa is any amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Pro Pro Pal Xaa Xaa Xaa Xaa
1 5

CLAIMS

1. A method for identifying a compound that binds to a target, the method comprising:

a) forming a first library comprising a multiplicity of peptides;

b) selecting from the first library at least one peptide that binds to the target;

c) determining the sequence or sequences of the at least one peptide that binds to the target, thereby generating a peptide motif;

d) forming a second library comprising a multiplicity of non-peptide compounds designed based on the peptide motif;

e) selecting from the second library at least one non-peptide compound that binds to the target; and

f) determining the structure or structures of the at least one non-peptide compound that binds to the target;

thereby identifying a compound that binds to the target.

2. The method of claim 1, wherein the first library is a phage display library.

3. The method of claim 1, wherein the first library is bound to a solid-support.

4. The method of claim 1, wherein the first library is an anchor library.

5. The method of claim 1, wherein the first library comprises at least about 10^6 peptides.

6. The method of claim 1, wherein the first library comprises at least about 10^9 peptides.

7. The method of claim 1, wherein the first library comprises at least about 10^{12} peptides.

8. The method of claim 1, wherein step c) comprises determining the nucleotide sequence of a nucleic acid molecule or molecules that encode the at least one peptide.

9. The method of claim 1, wherein step c) comprises determining the amino acid sequence or sequences of the at least one peptide.

10. The method of claim 1, wherein the second library comprises at least one peptide derivative.

11. The method of claim 1, wherein the second library comprises at least one peptide analogue.

12. The method of claim 1, wherein the second library comprises at least one peptidomimetic.

13. The method of claim 1, wherein the second library comprises at least about 10^2 non-peptide compounds.

14. The method of claim 1, wherein the second library comprises at least about 10^4 non-peptide compounds.

15. The method of claim 1, wherein the second library comprises at least about 10^6 non-peptide compounds.

16. The method of claim 1, wherein step f) comprises analyzing the at least one non-peptide compound by a mass spectrometric method.

17. The method of claim 16, wherein the mass spectrometric method comprises tandem mass spectrometry.

18. The method of claim 1, wherein the compound that binds to a target has a binding affinity for the target, expressed as an apparent K_d, EC_{50} or IC_{50}, of at least about 10^{-7} M.

19. The method of claim 1, wherein the compound that binds to a target has a binding affinity for the target, expressed as an apparent K_d, EC_{50} or IC_{50}, of at least about 10^{-8} M.

20. The method of claim 1, wherein the compound that binds to a target has a binding affinity for the target, expressed as an apparent K_d, EC_{50} or IC_{50}, of at least about 10^{-9} M.

21. The method of claim 1, wherein the at least one non-peptide compound that binds to the target as selected in step e) has at least a 10-fold higher affinity for the target than the at least one peptide that binds the target as selected in step b).

22. The method of claim 1, wherein the at least one non-peptide compound that binds to the target as selected in step e) has at least a 100-fold higher affinity for the target than the at least one peptide that binds the target as selected in step b).

23. The method of claim 1, wherein the at least one non-peptide compound that binds to the target as selected in step e) has at least a 1000-fold higher affinity for the target than the at least one peptide that binds the target as selected in step b).

24. The method of claim 1, further comprising:

g) forming a third library comprising a multiplicity of non-peptide compounds designed based on the structure or structures of the non-peptide compound or compounds determined in step f);

h) selecting from the third library at least one non-peptide compound that binds to the target; and

i) determining the structure or structures of the at least one non-peptide compound selected in step h);

thereby identifying a compound that binds to the target.

25. A method for identifying a compound that binds to a target, the method comprising:

a) forming a first library comprising a multiplicity of peptides displayed on the surface of a bacteriophage;

b) selecting from the first library at least one peptide that binds to the target;

c) determining the sequence or sequences of the at least one peptide that binds to the target, thereby generating a peptide motif;

d) forming a second library comprising a multiplicity of non-peptide compounds designed based on the peptide motif;

e) selecting from the second library at least one non-peptide compound that binds to the target; and

f) determining the structure or structures of the at least one non-peptide compound that binds to the target by tandem mass spectrometry;

thereby identifying a compound that binds to the target.

26. A method for identifying a compound that binds to a target, the method comprising:

a) forming a first library comprising an anchor library of a multiplicity of peptides;

b) selecting from the first library at least one peptide that binds to the target;

c) determining the sequence or sequences of the at least one peptide that binds to the target, thereby generating a peptide motif;

d) forming a second library comprising a multiplicity of non-peptide compounds designed based on the peptide motif;

e) selecting from the second library at least one non-peptide compound that binds to the target; and

f) determining the structure or structures of the at least one non-peptide compound that binds to the target by tandem mass spectrometry;

thereby identifying a compound that binds to the target.

27. A compound identified by the method of claim 1.

28. The compound of claim 27, which is a peptidomimetic.

29. The compound of claim 27, which binds to the target with a binding affinity, expressed as an apparent K_d, EC_{50} or IC_{50}, of at least about 10^{-7} M.

30. The compound of claim 27, which binds to the target with a binding affinity, expressed as an apparent K_d, EC_{50} or IC_{50}, of at least about 10^{-8} M.

31. The compound of claim 27, which binds to the target with a binding affinity, expressed as an apparent K_d, EC_{50} or IC_{50}, of at least about 10^{-9} M.

32. The compound of claim 27, which binds to the target with at least a 10-fold higher affinity than the at least one peptide that binds the target as selected in step b).

33. The compound of claim 27, which binds to the target with at least a 100-fold higher affinity than the at least one peptide that binds the target as selected in step b).

34. The compound of claim 27, which binds to the target with at least a 1000-fold higher affinity than the at least one peptide that binds the target as selected in step b).

35. A library comprising a multiplicity of non-peptide compounds designed based on a peptide motif, wherein the peptide motif is determined by selecting from a peptide library at least one peptide that binds to a target, determining the sequence or sequences of the at least one peptide that binds to the target and determining a peptide motif.

36. The library of claim 35, wherein the library comprises at least one peptidomimetic.

37. The library of claim 35, wherein the library comprises at least about 10^2 non-peptide compounds.

38. The library of claim 35, wherein the library comprises at least about 10^4 non-peptide compounds.

39. The library of claim 35, wherein the library comprises at least about 10^6 non-peptide compounds.

40. The library of claim 35, wherein the multiplicity of non-peptide compounds are attached to a solid support.

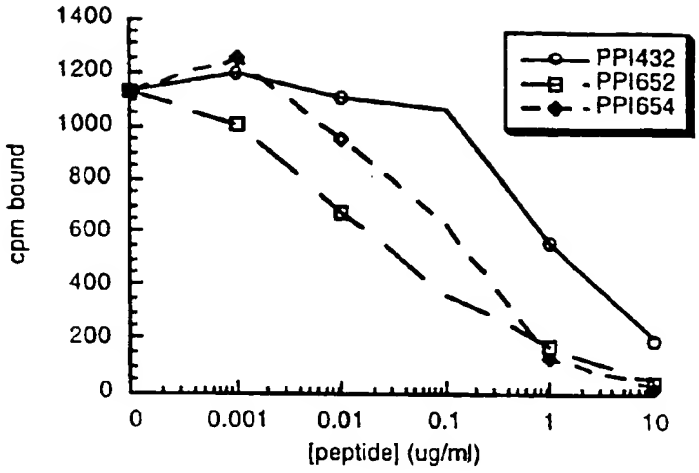

1400
1200
1000
800
600
400
200
0
cpm bound
0
0.001
0.01
0.1
1
10
[peptide] (ug/ml)
PPI432
PPI652
PPI654

FIGURE 1

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 96/20561

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C07K1/04 G01N33/68 C12Q1/68 C07K14/00

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C07K G01N C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0 639 584 A (INTERPHARM) 22 February 1995 see the whole document ---	1,3,5-40
X	WO 90 02809 A (PROTEIN ENGINEERING CORPORATION) 22 March 1990 see the whole document ---	2
A	WO 95 27072 A (PHARMAGENICS, INC.) 12 October 1995 see the whole document ---	1-40
A	WO 94 26775 A (HOUGHTEN PHARMACEUTICALS) 24 November 1994 see the whole document ---	1-40
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
7 April 1997	17.04.97

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+ 31-70) 340-2040, Tx. 31 651 epo nl, Fax (+ 31-70) 340-3016	Masturzo, P

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 96/20561

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF USA, vol. 87, 1990, WASHINGTON US, pages 6378-6382, XP000141872 S E CWIRLA ET AL.: "Peptides on phage: a vast library of peptides for identifying ligands" cited in the application see the whole document ---	2
E	WO 96 41180 A ((PHARMACEUTICAL PEPTIDES, INC.)) 19 December 1996 see the whole document -----	4

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter nal Application No
PCT/US 96/20561

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 639584 A	22-02-95	AU 6487394 A	05-01-95
		CA 2126359 A	23-12-94
		JP 7194382 A	01-08-95
		ZA 9404474 A	14-02-95
WO 9002809 A	22-03-90	AU 4308689 A	02-04-90
		EP 0436597 A	17-07-91
		JP 4502700 T	21-05-92
		US 5403484 A	04-04-95
		US 5571698 A	05-11-96
		US 5223409 A	29-06-93
WO 9527072 A	12-10-95	EP 0754238 A	22-01-97
WO 9426775 A	24-11-94	AU 6948294 A	12-12-94
		EP 0710245 A	08-05-96
WO 9641180 A	19-12-96	NONE	